



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/52, 9/00, C07K 16/40, C12N 15/11, 15/81, C12Q 1/25	A1	(11) International Publication Number: WO 99/32635 (43) International Publication Date: 1 July 1999 (01.07.99)
(21) International Application Number: PCT/GB98/03857 (22) International Filing Date: 21 December 1998 (21.12.98) (30) Priority Data: 9726897.3 20 December 1997 (20.12.97) GB (71) Applicant (for all designated States except US): ZENECA LIMITED [GB/GB]; 15 Stanhope Gate, London W1Y 6LN (GB). (72) Inventors; and (75) Inventors/Applicants (for US only): SCHNELL, Norbert, Friedemann [DE/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). DIXON, Graham, Keith [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). CHAVDA, Suberna [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). THAIN, John, Leslie [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). VINCENT, John, Philip [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). (74) Agent: PHILLIPS, Neil, Godfrey, Alasdair, Intellectual Property Dept., Zeneca Pharmaceuticals, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG (GB).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: ACETYL-COA-CARBOXYLASE FROM CANDIDA ALBICANS (57) Abstract The Acetyl-COA-carboxylase (ACCase) gene from <i>Candida albicans</i> .		

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

ACETYL-COA-CARBOXYLASE FROM CANDIDA ALBICANS

The present invention relates to Acetyl-CoA-carboxylase (ACCase) genes from *Candida Albicans* (*C. albicans*) and methods for its expression. The invention also relates to
5 novel hybrid organisms for use in such expression methods.

C. albicans is an important fungal pathogen and the most prominent target organism for antifungal research. ACCase is an enzyme of fatty acid biosynthesis and essential for fungal growth and viability. Inhibitors of the ACCase enzyme should therefore be potent antifungals. The ACCase proteins in all organisms are homologous to each other but they also differ
10 significantly in the amino acid sequence. Because selectivity problems (for example fungal versus human) it is extremely important to optimise potential inhibitor leads directly against the target enzyme (*C. albicans*) and not against a homologous but non-identical model protein, for example from *Saccharomyces cerevisiae* (*S. Cerevisiae*).

We have now successfully cloned the ACCase gene from *C. albicans* (hereinafter
15 referred to as the *C. Albicans* ACC1 gene) and elucidated its full length DNA sequence and corresponding polypeptide sequence, as set out in Figures 4 and 5 of this application respectively. The coding DNA sequence of the *C. Albicans* ACC1 gene is 6810 nucleotides in length and the corresponding protein sequence is 2270 amino acids in length. As will be explained below there are two forms of the *C. Albicans* ACC1 gene, the above numbers relate
20 to the longer version, Met1.

Therefore in a first aspect of the present invention we provide a polynucleotide encoding a *C. albicans* ACCase gene, in particular the (purified) *C. albicans* ACC1 gene as set out in Figure 5 hereinafter. It will be appreciated that the polynucleotide may comprise any of the degenerate codes for a particular amino acid including the use of rare codons. The
25 polynucleotide is conveniently as set out in Figure 4. It will be apparent from Figure 4 that the gene is characterised by the start codons Met1 and Met2 (as indicated by the first and second underlined atg codons, hereinafter referred to as atg1 and atg2 respectively). Both forms of the gene starting from Met1 and Met2 respectively are comprised in the present invention. The invention further comprises convenient fragments of any one of the above sequences.

Convenient fragments may be defined by restriction endonuclease digests of sequence, suitable fragments include a full length *C. Albicans* ACC1 gene (starting with Met1 or Met2) flanked by unique *StuI* (5'-end)-*NotI* (3'-end) restriction sites as detailed in Figure 6.

We also provide a polynucleotide probe comprising any one of the above sequences or
5 fragments together with a convenient label or marker, preferably a non-radioactive label or marker. Following procedures well known in the art, the probe may be used to identify corresponding nucleic acid sequences. Such sequences may be comprised in libraries, such as cDNA libraries. We also provide RNA transcripts corresponding to any of the above *C. Albicans* ACC1 sequences or fragments.

10 In a further aspect of the invention we provide a *C. albicans* ACC1 enzyme, especially the ACC1 enzyme having the polypeptide sequence set out in Figure 5, in isolated and purified form. This is conveniently achieved by expression of the coding DNA sequence of the *C. Albicans* ACC1 gene set out in Figure 4, using methods well known in the art (for example as described in the Maniatis cloning manual - Molecular Cloning: A Laboratory Manual, 2nd
15 Edition 1989, J. Sambrook, E.F. Fritsch & Maniatis). As indicated for Figure 4 above, the enzyme is characterised by two forms Met1 and Met2. Both form of the enzyme are comprised in the present invention.

The *C. Albicans* ACC1 enzyme of the present invention is useful as a target in biochemical assays. However, to provide sufficient enzyme for a biochemical assay for *C. Albicans* ACC1 (for example, for a high throughput screen for enzyme inhibitors) this has to be
20 purified. Two major constraints impair this purification.

- 1) any new organism will necessitate deviation from published procedures because it will differ in its lysis and protease activity. *C. albicans* is known to express and secrete many aspartyl proteases.
- 25 2) The expression of *C. Albicans* ACC1 is very low and satisfying purification results can only be achieved if the enzyme is overexpressed.

We have now been able to overcome these problems by controlled overexpression of the *C. albicans* ACC1 in a *Saccharomyces* strain. This means that subsequent purification of the enzyme may then for example follow published procedures.

Therefore in a further aspect of the present invention we provide a novel expression system for expression of a *C. albicans* ACC1 gene which system comprises an *S. cerevisiae* host strain having a *C. albicans* ACC1 gene inserted in place of the native ACC1 gene from *S. Cerevisiae*, whereby the *C. albicans* ACC1 gene is expressed. Preferred *S. cerevisiae* strains
5 include JK9-3D α and its haploid segregants.

The *C. albicans* ACC1 gene is preferably over-expressed relative to that as may be achieved by a *C. albicans* wild type strain, ie under the control of its own ACC1 promoter. Whilst we do not wish to be bound by theoretical considerations, we have achieved approximately 14 fold over-expression relative to the wild-type host *S. cerevisiae* strain JK9-3D.
10 This may be achieved by replacing the *C. albicans* promoter in the expression construct by a stronger and preferably inducible promoter such as the *S. cerevisiae* GAL1 promoter.

Controlled overexpression is used to improve expression of a *C. albicans* polypeptide relative to expression under the control of a *C. albicans* promoter. In addition using procedures outlined in the accompanying examples we have been able to isolate a fully functional *C.*
15 *albicans* ACC1 gene as determined by 100% inhibition by SoraphenA.

The novel expression system is conveniently prepared by transformation of a heterozygous ACC1 deletion strain of a convenient *S. cerevisiae* host by a convenient plasmid comprising the *C. albicans* ACC1 gene. Transformation is conveniently effected using methods well known in the art of molecular biology (Ito et al. 1983).

20 The plasmid comprising the *C. albicans* ACC1 gene and used to transform a convenient *S. cerevisiae* host represents a further aspect of the invention. Preferred plasmids for insertion of the *C. Albicans* ACC1 gene include YEp24, pRS316 and pYES2(Invitrogen).

The heterozygous ACC1 deletion strain of a convenient (diploid) *S. cerevisiae* host is conveniently achieved by disruption preferably using an antibiotic resistance cassette such as
25 the kanamycin resistance cassette described by Wach et al (Yeast, 1994, 10, 1793-1808).

The expression systems of the invention may be used together with, for example cell growth and enzyme isolation procedures identical to or analogous to those described herein, to provide an acetyl-COA-carboxylase (ACCase) gene from *C. albicans* in sufficient quantity and with sufficient activity for compound screening purposes.

In a further aspect of the invention we provide the use of an acetyl-CoA-carboxylase (ACCase) gene from *C.albicans* in assays to identify inhibitors of the polypeptide. In particular we provide the their use in pharmaceutical or agrochemical research.

As presented above the *C. albicans* ACC1 enzyme may be used in biochemical assays to
5 identify agents which modulate the activity of the enzyme. The design and implementation of such assays will be evident to the biochemist of ordinary skill. The enzyme may be used to turn over a convenient substrate whilst incorporating/losing a labelled component to define a test system. Test compounds are then introduced into the test system and measurements made to determine their effect on enzyme activity. Particular assays are those used to identify
10 inhibitors of the enzyme useful as antifungal agents. By way of non-limiting example, the activity of the ACC1 enzyme may be determined by (i) following the incorporation (HCO_3 , Acetyl-CoA) or loss (ATP) of a convenient label from the relevant substrate (T.Tanabe et al, Methods in Enzymology, 1981, 71, 5-60; M. Matasuhashi, Methods in Enzymology, 1969, 14, 3-16), (ii) following the release of inorganic phosphate from ATP (P. Lanzetta et al, Anal.
15 Biochem. 1979, 100, 95-97), or (iii) following the oxidation of NADH in a coupled assay, for example using either fatty acid synthetase or pyruvate kinase/lactate dehydrogenase enzymes. Convenient labels include carbon14, tritium, phosphorous32 or 33.

Any convenient test compound(s) or library of test compounds may be used. Particular test compounds include low molecular weight chemical compounds (molecular weight less than
20 1500 daltons) suitable as pharmaceutical agents for human, animal or plant use.

The enzyme of the invention, and convenient fragments thereof may be used to raise antibodies. Such antibodies have a number of uses which will be evident to the molecular biologist of ordinary skill. Such uses include (i) monitoring enzyme expression, (ii) the development of assays to measure enzyme activity and precipitation of the enzyme.

25 In addition we provide antisense polynucleotides specific for all or a part of an ACC1 polynucleotide of the invention.

The invention will now be illustrated but not limited by reference to the following Table, Example, References and Figures wherein:

Table 1 shows the comparative properties of native and recombinant acetyl-CoA carboxylase enzymes

Figure 1 shows partial sequence from the *C. albicans* genome. Underlined regions were used to derive PCR primers, to generate a *C. albicans* ACC1 specific probe.

Figure 2 shows cloned fragments of the *C. albicans* ACC1 gene isolated from genomic DNA libraries. Arrows indicate extension of the fragment beyond the region displayed.

Figure 3 shows sequenced XbaI-HinDIII and HinDIII subclones of clone CLS1-b1.

Figure 4 shows the full DNA sequence of the *C. albicans* ACC1 gene. The atg start codons for Met1 and Met 2 are in lower case and underlined, as is the tag stop codon.

Figure 5 shows the full protein sequence of the *C. albicans* ACC1 gene. Putative start codons for Met1 and Met2 are shown in bold.

Figure 6 shows the generation of a tailored ACC1 gene (minus promoter) for expression under control of the GAL1 promoter in plasmid pYES2. From the initial ACCase gene (line1) the core SacI-BamHI (line3) is modified by the addition of 3' BamHI-NotI (line2) and 5' StuI-SacI (different fragments for Met1 and -2 lines 5 and 7 respectively) to generate the final "portable" gene flanked by StuI-NotI (lines 6 and 8).

Figure 7 shows the results of the *in-vitro* ACCase enzyme assay set out in the accompanying Example when Soraphen A (a specific inhibitor of the ACCase enzyme) was supplied (X-axis) over the range 0.1nM-100µM in the dose response regimen of the assay.

Example 1

Cloning of the *C. albicans* ACC1 gene and generation of a heterologous *S. cerevisiae* expression system:

1) Probe generation

We used the polymerase chain reaction (PCR) to generate a DNA probe between and including the underlined regions in Figure 1

2) Identification of clones from a *C. albicans* genomic library hybridising to the ACCase probe

The PCR product was labelled using an "ECL direct nucleic acid labelling and detection kit" (Amersham) as described by the supplier. The PCR product (probe) was then
5 shown to hybridise to *S. cerevisiae* (weakly) and *C. albicans* genomic DNA. in a Southern blot procedure (as described Maniatis, 1989). Two genomic DNA libraries (CLS1 and CLS2) of *C. albicans* (in the yeast-*E. coli* shuttle plasmids YEp24 and pRS316 respectively, (as described in Sherlock et al. 1994, source: Prof. John Rosamond, Manchester University) were used to isolate fragments hybridising with the probe which was radiolabelled using "Ready To
10 Go" dCTP labelling beads (Pharmacia, as described by the manufacturer). The colony hybridisation was carried out as described by Maniatis (1989). Hybridising colonies were identified, plasmid DNA isolated, purified (Quiagen maxiprep, as described by the supplier) and sequenced (Applied Biosystems, model 377 sequencer) from their junctions with the plasmid. Several fragments carrying partial ACCase gene sequence as well as one full length
15 clone could be identified (Figure 2).

3) Sequencing of the cloned gene, comparison with ACCases from *S. cerevisiae*, other fungi and higher eukaryotes (plants, mammals, man)

The bulk of the sequence of the *C. albicans* ACC1 gene was determined (on both
20 strands) using flanking sequence- or insert sequence-specific primers from defined *Hin*DIII and *Xba*I-*Hin*DIII subfragments (of clone CLS1-b1) cloned into pUC19 (see Figure 4). The promoter and 5' coding region absent from this clone was established from CLS2-d1 and the gene's 3' end from CLS2-13 using insert specific primers. All junctions including the ones between the *Hin*DIII subfragments were verified from the full length clone CLS2-13 (in
25 YEp24. The full length DNA sequence of *C. albicans* (Ca) ACC1 is shown in Figure 5a and the protein translation in Figure 5b. The two potential start Methionines, Met1 and Met2 are shown in bold

The protein is homologous to ACCases of other fungi (*S. cerevisiae*, *S. pombe* and

U maydis) and also to the plant (*Brassica napus*), mammalian (sheep, chicken and rat) and human enzymes. Of the two potential start codons of *C. albicans* ACC1, Met 2 seems the more likely one as the sequence between Met1 and Met2 is unrelated to the other ACCases and indeed to any other protein sequence in the EMBL/Genbank database. The high degree of
5 homology between ACCases of different species and the apparent lack of an identifiable fungal subgroup makes it even more important to use the actual target enzyme (here from the pathogen *C. albicans*) as a screening tool to identify specific inhibitors.

4) Generation of a heterozygous ACC1 deletion strain of *S. cerevisiae*

10 As ACCase is an essential enzyme, only one allele of a diploid cell can be deleted without loss of survival. One ACC1 gene of a diploid *S. cerevisiae* strain (JK9-3Daa, Kunz et al. 1993) was therefore disrupted using the kanamycin resistance cassette as described by Wach et al. using the protocol described therein. Sporulation of the heterozygous diploid (ACC1/acc1::KANMx) yields only two viable spores (which are kanamycin-sensitive)
15 showing the essentiality of the ACC1 gene as well as the characteristic arrest phenotype for the two inviable spores (as published by Haßbacher et al., 1993).

5) Complementation of a *S. cerevisiae* ACC1 deletion with the cloned *Candida* gene, Ca ACC1

20 The heterozygous ACC1/acc1::KANMx strain was transformed with one full length *C. albicans* gene (CLS2-13 in Yep24). Expression of the gene from this plasmid will be due to functionality of the *Candida* ACC1 promoter in the heterologous *S. cerevisiae* system. Complementation of the knockout was demonstrated by sporulating the diploid transformants. In most cases 3-4 viable (haploid) spores were detected. The analysis of tetrads indicated that
25 kanamycin-resistant colonies were only formed if they also contained the complementing CLS2-13 plasmid, as indicated by the presence of the URA3 transformation marker. This clearly shows that the *C. albicans* gene fully complements the ACCase function in *S. cerevisiae*. Therefore the strain generated can be used to screen for inhibitors which are specific for the *Candida* enzyme in the absence of a background of *Saccharomyces* enzyme.

As demonstrated by its functionality, the heterologous protein folds correctly in the host, *S. cerevisiae*, where it must also have been correctly biotinylated by the *S. cerevisiae* machinery (carried out by ACC2, encoding protein-biotin-ligase).

To facilitate purification of *C. albicans* ACCase, it is beneficial to achieve
5 overexpression of the protein in a suitable host. Therefore the *C. albicans* promoter was replaced by the stronger and inducible *S. cerevisiae* GAL1 promoter. As the *Candida* sequence had revealed two potential start codons (see Figure 4) for the ACC1 reading frame, both versions were placed under GAL1 control. To generate appropriate restriction sites for cloning, the ACC1 gene was modified via PCR at both ends (see Figure 6 above). and cloned
10 into plasmid pYES2 (Invitrogen) as a *Stu*I-*Not*I fragment into *Hin*DIII (fill-in)-*Not*I sites of the vector. The identity of the PCR-modified gene-parts with the original ones was confirmed by sequencing. Both constructs (Met1 and Met2) complement the *S. cerevisiae* ACC1 knockout when the cells are grown on galactose but not on glucose (where the GAL1 promoter is switched off). Growth is very poor if the gene is transcribed initiating at Met1,
15 whereas Met2 restores wild type growth rates in *S. cerevisiae*.

6) Overexpression of the *Ca* ACCase to facilitate protein purification and use for screening purposes

20 Materials

Growth Media :-

Sabouraud Dextrose broth

Yeast peptone dextrose broth (YPD)

Yeast peptone galactose broth (YPGal) (i.e. 2% w/v galactose)

25

Growth of cells

Candida albicans B2630 (Janssen Pharmaceutica, Beerse, Belgium) was maintained on Sabouraud dextrose agar slopes at 37 °C which were subcultured biweekly. For the growth of liquid cultures for experiments, *C. albicans* grown on Sabourauds dextrose agar for

48 h at 37°C was used to inoculate 50 ml Sabouraud dextrose broth containing 500µg/l d-biotin. This was incubated for 16 h at 37 °C on a platform shaker (150 rpm). 1.5 ml of this culture was added to each of 24 x2 litre conical flasks, each containing 1 litre of Sabouraud dextrose broth containing 500µg/l d-biotin, giving a final inoculum concentration of

5 approximately 1.5×10^6 cfu ml⁻¹. The cultures were grown for 9 h, at 37 °C (log phase) with shaking (150 rpm). Cell numbers in liquid culture were determined spectrophotometrically (Philips PU8630 UV/VIS/NIR Spectrophotometer) at 540 nm in a 1 cm path length cuvette. Absorbance was linearly related to cell number up to an OD. of 2.0.

Saccharomyces cerevisiae strains Mey134 and CLS2-13 were maintained on Yeast

10 peptone dextrose (YPD) agar plates at 30 °C, which were subcultured biweekly. For the growth of liquid cultures for experiments, the *S. cerevisiae* strains were grown on YPD agar for 48 h at 30 °C and were then used to inoculate 50 ml YPD broth containing 500µg/l d-biotin, which was incubated at 30°C for 16h on a platform shaker (200 rpm). 2.0 ml of this culture (approx. 4×10^8 cfu/ml) was added to each of 24 x 2 litre conical flasks, each

15 containing 1 litre of YPD broth containing 500µg/l d-biotin, giving a final inoculum concentration of approximately 8×10^5 cfu/ml. The cultures were grown for 9 h, at 30 °C (log phase) with shaking (200 rpm). Cell numbers in liquid culture were determined spectrophotometrically (Philips PU8630 UV/VIS/NIR Spectrophotometer) at 540 nm in a 1 cm path length cuvette.

20 *Saccharomyces cerevisiae* strains PNS117a 5C, PNS117b 6A, and PNS 120a 6C were maintained on Yeast peptone galactose (YPGal) agar plates at 30 °C which were subcultured biweekly. For the growth of liquid cultures for experiments, the *S. cerevisiae* strains were grown on YPGal agar for 48 h at 30 °C and were then used to inoculate 50 ml YPGal broth containing 500µg/l d-biotin and 200µg/ml kanomycin, which were incubated at 30°C for 30h

25 on a platform shaker (200 rpm). 2.0 ml of this culture (approx. 4×10^8 cfu/ml) was added to each of 24 x2 litre conical flasks, each containing 1 litre of YPGal broth containing 500µg/l d-biotin and 200µg/ml kanomycin, giving a final inoculum concentration of approximately 8

x10⁵ cfu/ml. The cultures were grown for approximately 23h at 30 °C (log phase) with shaking (200 rpm).

Determination of cell number

- 5 Cell numbers were determined using a standard viable count agar based plating method, using the appropriate agar media.

Preparation of fungal ACCase enzyme

- Cultures of the appropriate yeast strains were grown to the exponential phase of growth (for *Saccharomyces* and *Candida* strains respectively). These were then harvested by centrifugation (4400 g, 10min, 4 °C), washed twice in 700ml of 50mM Tris pH7.5 containing 20% w/v glycerol, resuspending the cell pellet each time. The final washed pellet was fully resuspended into a thick slurry using 10 to 20ml of buffer (50mM Tris pH7.5 containing 1mM EGTA, 1mM EDTA (disodium salt), 1mM DTT, 0.25mM Pefabloc hydrochloride, 1µM Leupeptin hemisulphate, 1µM Pepstatin A, 0.5µM Trypsin inhibitor and 20% w/v glycerol). The volume of buffer required was dependent on the total packed cell wet weight. (i.e. 1ml buffer added per 6gm of packed wet cell pellet).
- 10
15

- The cell paste was homogenised using a pre-cooled Bead-Beater (Biospec Products, Bartlesville, OK 74005) with 4 x 10 second Bursts, allowing 20 second intervals on ice. The preparation was then centrifuged at 31,180g for 30 minutes. After centrifugation the supernatant was immediately decanted into a container, then aliquoted before snap freezing in liquid nitrogen. The preparation was then stored at -80°C and was found to be stable for at least 2 months.
- 20

- All enzyme preparation steps were carried out at +4°C, unless otherwise stated.
- 25

In-vitro ACCase enzyme assay

The assay was conducted in 96 well, flat bottomed polystyrene microtitre plates. All test and control samples were tested in duplicate in this assay.

100µl of the ACCase enzyme preparation (in 50mM Tris pH7.5 containing 1mM EGTA, 1mM EDTA (disodium salt), 1mM DTT, 0.25mM Pefabloc hydrochloride, 1µM Leupeptin hemisulphate, 1µM Pepstatin A, 0.5µM Trypsin inhibitor, and 20% w/v glycerol) was added to each well of the microtitre plate. Each well contained either a 3µl test sample
5 made up in DMSO or 3µl DMSO alone (NB. Final DMSO concentrations in the assay were 1.48% v/v). The microtitre plates were placed in a water bath maintained at 37°C. 10µl of [¹⁴C] NaHCO₃ containing 9.25kBq in 378mM NaHCO₃ was then added to each well. The reaction was initiated by the addition of 100µl of Acetyl Coenzyme A containing assay buffer (50mM Tris pH7.5 containing 4.41mM ATP(disodium salt), 2.1mM Acetyl
10 Coenzyme A, 2.52mM DTT, 10.5mM MgCl₂, and 0.21% w/v Albumin [Bovine, fraction V]), (removed from ice 5 minutes before use) to each well. The tubes were incubated at 37°C for 5 minutes. The reaction was then terminated by the addition of 50µl of 6M HCl to each well. In parallel, a pre-stopped assay control was set up which involved adding the 50µl of 6M HCl prior to [¹⁴C] NaHCO₃ and the assay buffer (No further HCl additions were made to
15 these wells after the 5 minute incubation). The DPM values for the pre-stopped assay were subtracted from the normal assay situation.

After the addition of the stop reagent the plates were left open in the water bath for a further 30 minutes to allow the ¹⁴CO₂ to escape. After this time 150µl of each reaction mixture were applied onto individual GF/C glass microfibre filter discs and allowed to dry
20 thoroughly before adding scintillation fluid. Radioactivity in the samples was then determined by scintillation counting (Wallac WinSpectral 1414, Turku, Finland).

IC50's were calculated from the data using non-linear regression techniques available in the ORIGIN software package (Microcal Software Inc., Massachusetts, USA).

Soraphen A which is a specific inhibitor of ACCase was supplied over the range
25 0.1nM-100µM in the dose response regimen of the assay.

Protein determination

The total protein concentration of each ACCase preparation used was determined by the Coomassie Blue method (Pierce, Illinois, USA), (using 1cm path length cuvettes read 595nm (Philips PU8630 UV/VIS/NIR Spectrophotometer).

5

In-vitro antifungal activity

Compounds were tested over a concentration range of 1024 - 0.00098 µg/ml by a broth-dilution method in microtitre plates using doubling dilutions in YPD or YPGal (both containing 500µg/l d-biotin). Stock solutions of inhibitors were prepared at 51.2mg/ml in
10 Dimethyl sulphoxide (DMSO) (final assay concentration of DMSO was 2% v/v). Each Yeast culture was added to the well to give a final 10⁴ cfu/well. The plates were incubated at 30°C for 48h and MIC's determined visually.

Discussion

15 Expression of ACCase, a biotinylated protein, was monitored by a "biotin-avidin affinity western blot" as described by Haßlacher et al., 1993. Expression of the *C. albicans* ACC1 gene from its own promoter from plasmid Yep24 was comparable to that of the *S. cerevisiae* gene (no overexpression). Expression under control of the GAL1 promoter however, was considerably higher indicating a drastically increased level of biotinylated and
20 therefore fully functional enzyme. Transcription of the gene was fully induced as the cells had to be grown on galactose to be viable. On glucose the GAL1 promoter is completely off, causing the cells to arrest and eventually die due to insufficient supply of ACCase). The *S. cerevisiae* strain described in this application is a convenient source of the *C. albicans* enzyme. The engineered strain possesses no residual background ACCase because the gene
25 coding for the *S. cerevisiae* enzyme had been removed. Congenic versions of such a strain (genetically identical apart from the ACCase gene carried) expressing different ACCases (e.g. the different human (Abu-Elheiga et al. 1995), mammalian (Lopez-Casillas et al., 1988, Takai et al. 1988, Barber et al., 1995)), plant (Schulte et al., 1994) or other fungal enzymes (Al-Feel et al., 1992, Saito et al., 1996, Bailey et al., 1995)) can be used as tools for

screening. Differences in growth of such strains may be solely dependent on differences in their ACCase activity. Differential growth in the presence of ACCase inhibitors (for example soraphenA or compounds yet to be identified) indicates selectivity of the drug towards one type of the ACCase enzyme.

5

References:

- Abu-Elheiga L., Jayakumar A., Baldini A., Chirala S.S., Wakil S.J.; Proc. Natl. Acad. Sci. U.S.A. 92: 4011-4015(1995).
- Al-Feel W., Chirala S.S., Wakil S.J.; Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538(1992).
- 10 Bailey A.M., Keon J.P.R., Owen J., Hargreaves J.A.; Mol. Gen. Genet. 249:191-201(1995).
- Barber M.C., Travers M.T.; Gene 154:271-275(1995).
- Haßlacher M., Ivessa A. S., Paltauf F., Kohlwein S. D.; J. Biol. Chem. 268:10946-10952 (1988).
- Ito, H., Fukuda, Y., Murata, K., Kimura, A.; J. Bacteriol. 153: 163-168 (1983)
- 15 Kunz, J., Henriquez, R., Schneider, U., Deuter-Reinhard, M., Movva, N.R., Hall, M.N.; Cell 73: 585-596 (1993)
- Lopez-Casillas F., Bai D.-H., Luo X., Kong I.-S., Hermodson M.A., Kim K.-H.; Proc. Natl. Acad. Sci. U.S.A. 85:5784-5788(1988).
- Maniatis T., Frisch E. F., Sambrook J.; Molecular Cloning, Cold Spring Harbour Laboratory
- 20 Press (1989)
- Saiki R. K., Gelfand D. H., Stoffel S., Sharf S. J., Higuchi R., Horn G. T., Mullis K. B., Erlich H. A.; Science 239: 487-494 (1988)
- Saito A., Kazuta Y., Toh H., Kondo H., Tanabe T.; S. pombe ACC1, Submitted (Dec-1996) to Embl/Genbank/Ddbj Data Banks.
- 25 Schulte W., Schell J., Toepfer R.; Plant Physiol. 106:793-794(1994).
- Sherlock G., Bahman A. M., Mahal A., Shieh J. C., Fewtrell M., Rosamond J.; Mol. Gen. Genet. 245: 716-723.
- Takai T., Yokoyama C., Wada K., Tanabe T.; J. Biol. Chem. 263:2651-2657(1988).
- Wach A., Brachat A., Pöhlmann R., Philippsen P.; Yeast 10: 1793-1808 (1994)

TABLE I
Comparative properties of native and recombinant acetyl-CoA carboxylase enzymes

Yeast strain	Cell doubling time (minutes)	Growth temperature for ACCase preparation (°C)	Liquid MIC (µg/ml) for Soraphen A	IC50 for Soraphen A (nM) against ACCase preparations	Specific activity of ACCase preparation (nmoles product/min/mg protein)
<i>C. albicans</i> B2630	56	37	0.003		
<i>S. cerevisiae</i> Mey 134	160	30	8		0.641
<i>S. cerevisiae</i> CLS2-13	163	30	2	2.499	3.054
<i>S. cerevisiae</i> PNS 117a 5C	253	30	2	17.518	7.025
<i>S. cerevisiae</i> PNS 117b 6A	222	30	4	13.083	10.573
<i>S. cerevisiae</i> PNS 120a 6C	303	30	0.5	ND	0.244
<i>S. cerevisiae</i> PNS 120b 1C	287	30	0.125	ND	ND

Key :- ND =
not determined

Claims:

1. A polynucleotide encoding an Acetyl-COA-carboxylase (ACCase) gene from *Candida albicans*.

5

2. A polynucleotide as claimed in claim 1 and as set out in Figure 4 herein.

3. A polynucleotide as claimed in claim 2 and characterised by the start codon atg2.

10 4. A polynucleotide comprising a restriction fragment of a polynucleotide as claimed in any one of claims 1-3.

5. A polynucleotide probe comprising a polynucleotide as claimed in any one of claims 1-4.

15

6. An Acetyl-COA-carboxylase (ACCase) polypeptide from *Candida albicans* in isolated and purified form.

7. A polypeptide as claimed in claim 6 and as set out in Figure 5.

20

8. A polypeptide as claimed in claim 7 and characterised by Met2.

9. A polypeptide as claimed in claim 6 and obtained by expression of a polynucleotide as claimed in any one of claims 1-4.

25

10. Antibodies specific for a polypeptide as claimed in any one of claims 6-9.

11. An antisense polynucleotide specific for all or a part of a polynucleotide as claimed in any one of claims 1-4.

30

12. An RNA transcript corresponding to a polynucleotide as claimed in any one of claims 1-4.

13. An expression system for expression of an Acetyl-COA-carboxylase (ACCase)
5 polypeptide from *Candida albicans* which system comprises an *S. cerevisiae* host strain having a *Candida albicans* ACC1 polynucleotide as claimed in any one of claims 1-3, inserted in place of the native ACC1 gene from *S. Cerevisiae*, whereby the *Candida albicans* ACC1 polypeptide is expressed.

10 14. An expression system as claimed in claim 13 and adapted for controlled overexpression of the *Candida albicans* polynucleotide relative to expression under the control of a *Candida albicans* promoter

15. An expression system as claimed in claim 14 and used to provide an Acetyl-COA-
15 carboxylase (ACCase) gene from *Candida albicans* in sufficient quantity and with sufficient activity for compound screening purposes.

16. Use of an Acetyl-COA-carboxylase (ACCase) polypeptide from *Candida albicans* as claimed in claim 6, in an assay to identify inhibitors of the polypeptide.

20

17. Use as claimed in claim 16 in pharmaceutical research.

FIGURE 1

5

10

GCACGCTTGACGGTTTTACCAAATGCGAAAATATGACCAAATTGAGAATCCGAAAATGA
ATGGATAGAAGATTGGTTACCAACTGAGAAAATAACCCACACATTAGAAGAAGAACGGAA
ATTCAATTCATGTAAAGAACCACCACTTGGTTTAAAACCTTCACCAGGATCTTCAGAAGT
AATACGACAAGCAGTACAATGTCCCTTTGGTGTGGTCTTCTTTGACTAACCAATGAAGT
TTCTGACTTGAATTCAAAATCAATATCAGTAGTGGTATGAGGATCGGCACCGCACAAAGT
TCTGATATCTCTGATTCTATGCATTGGTATACCCATAGCAATTTGTAATTGAGCAGCTGG
TAAATTAACACCTGTCACCATTTCACTGGTTGGATGTTCAACTTGCAATCTTGGGTTCAA
TTCCAAAAAGTAGAATTTATCTTCAGCGTGGGGAGTAAAGGTACTCAACAGTACCAGGGG
GTTACATAACCAACTTATTTTACCCAATCTGACTGGTGGATT

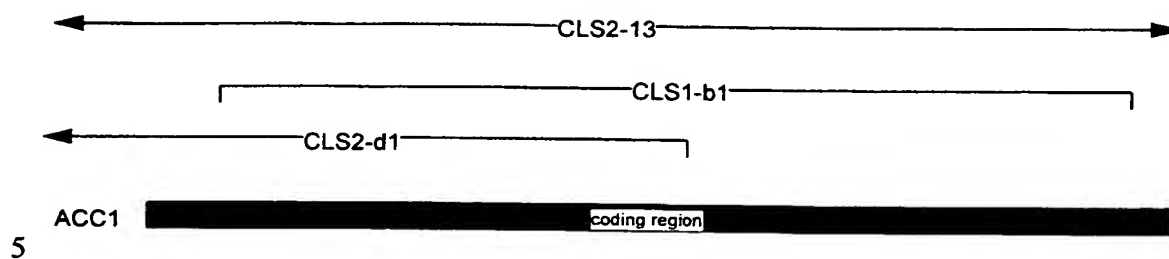
FIGURE 2

FIGURE 3

5

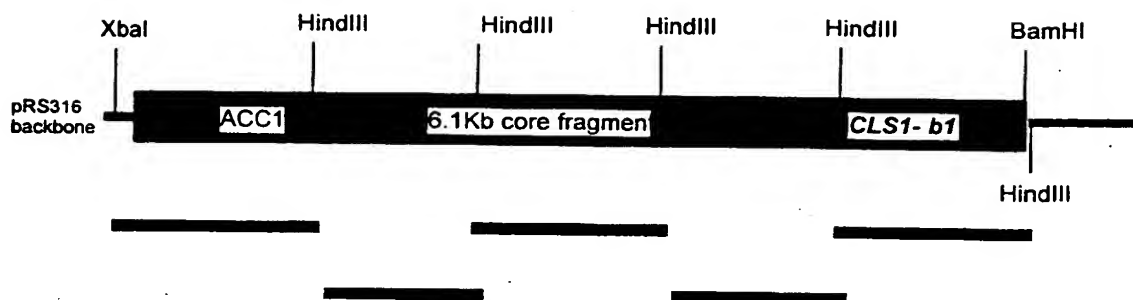


FIGURE 4

AATATATTGCTTCCTTTTGATAGGAAGTAACTCCGAGTGTTTGAATTTGATATATGTTATTTCATATACGTTCAATGGCTC
TCTTCTATGCTTTGTATATACTTTCTTTTGAATAGATACTCATGTAAAGAGATTTGAAACCATATTCTAACCAACAAAA
5 TATTGTACGGTATAGGTTAGAAAAAACTCCGTAAGGTCGCTTACACGGTTAAATTGAAAACACGTTAAAAATATATT
TGGGTAATGGACTAAGCTATATACAGTACTCAACAAAAATGAAATCAACACAATGTTCTTTGGGAAATTCATTTTCATGC
AATAGGGTGATTCTCTTTCTACTATCCAACAACGATACCCCTGCTTTTGAATAATCTTTCTAAATTCAAATTTGATATA
ATTCTTATTTATATATTACTTTCTTTTCCCATATAACCCCATTTTTTTTTTGGAAATCATATTTGTTTGGATTGGCT
10 TTCCCTTTTCAGTCTGAGGAACATACTAATTACGAACAACAATTATACATCCAATCTTCATCTAACGAATTGATTATTTAC
ATTTATTAACCCCTTGGATACAACTGATTACACTTTTTAGTTAGTTTGTTCATTTATAAGGGTATTATACAACAAAGAT
ATCATTTAAAGTTAAATCTCAATCTGGAATAATAAAAGTATTCAACACTTTTGTCTACAATAGGTATGTTCAAAATCAAT
TGAAGCCATCGAGATAAGAAATTAAGCAAAAACGTTTACATTTGTTGTGTGTGTGTGTCAGTGTGGAAGAAGCTCGAGT
GATTGCTTTTCTCGGCATCAGCTGTGTTGGGAACATCTTGTGCTTAAAGTTTCGGAGTAATATTAGAGTAATGGAACGA
AAAAACAAAATAAAGTTCTGGAACCAAAAGATTTGAAAAATTGGGTAGAAAACAAAAAAGACAAAGCAGGAACCCAA
15 CAATAAATGAACTAACTCAAAACTACTCACAACAACAACACTTATTTTCACTTGCTTTATTTCTTCGATTTTTTatg
AGATGCAATTTATCTCTAATAAAGAATACTAATCACTTGCTTACATGATCGGCTTCCCTAATTACAAAACCAACTATA
TATACCTCATCTCATTATATCCCATTCAGAACATATTCAAGTCATTGTTAatgTCAGATCAATCTCCATCTCCCTAGTC
CTAGCGATTCCCTTAGCTACACTACATTACATGAAAATTGGCCATCTCATTCTTGGGTGGAAATTCAGTTTGAATGCT
GAACCTTTAAAGTCAGAGACTTTGTGAGAGCTCATCAAGGTCATACAGTTATTTTCGAAAATTTAATTGCCAACATGG
20 TATAGCTGCACTTAAAGAAATCAGATCAGTTAGAAAAATGGCTTATGAAACATTTGGTGACGAAAAGCCATACAGTTTA
CCGTTATGGGCACTCCAGAAGATTGGAAGCTAATGCCGAATATATTAGAATGGCCGACCAATTCATTGAAGTCCCTGGT
GGCACCATAACAATAACTATGCTAATGTTGATCTCATTTGATAGATAGCAGAAAGTACAATGCTCATGCCGTTTGGGC
TGGGTGGGGCATGCTTCAGAGAATCCTTTGTTACCAGAAAAATTAGCTGCATCTCCCAAAAAAATTATTTTATTTGGTC
CTCCTGTTTCAGCTATGAGATCTTTAGGTGACAAGATTTGCTTACTATAGTTGCTCAACATGCTCAAGTACCATGTATT
25 CCATGGTCCGGTACTGGTGTGATGAAGTGAATAAGACCCACAACTAATTTGGTTTCTGTTGCTGATGATTTATGTC
CAAAGGGTGCTGTACTAGTCCAGAAGATGGTTTAGAAAAAGCCAAAAAATTGGGTTCCAGTTATGATTAAAGCCTCTG
AAGTGGTGGTGTAAAGGTATTAGAAAAGTTGATGATGAGAAAACTTCATTACCTTATACAACCAAGCAGCTAATGAA
ATACCAGGTTCTCCTATCTTTATGAAAGTGTGATGAGAAAACTTCATTACCTTATACAACCAAGCAGCTAATGAA
CGGTACTAACATTTCCCTTTTTGGAAGAGATTGTTCCGTACAAAGAAGACACCAAAAGATTATTGAAGAAGCACCAGTCA
30 CCATTGCCAGAAAGGAACTTTCCACGAAATGGAAATGCAGCAGTCAGATTGGGTAAATTAGTTGGTTATGATCCGCT
GGTACTGTTGAGTATCTTTACTCCACGCTGAAGATAAATCTACTTTTTGGAATTGAACCAAGATTGCAAGTTGAACA
TCCAACCACTGAAATGGTGACAGGTGTTAATTTACAGCTGCTGCTTACTTACAAATTTGCTATGGGTATACCAATGCATAGAA
TCAGAGATATCAGAATTTGTACGGTGCCGATCCTCATACCCTACTGATATTGATTTTGAATTCAGTCAAGTCAAGTCA
TTGGTTACTCAAAGAAGACCAACACCAAGGGACATTTGACTGCTTGTGCTATTACTTCTGAAGATCCTGGTGAAGGTTT
35 TAAACCAAGTGGTGCTTTTACATGAATTTCCGTTCTTCTTCTAATGTGTGGGGTTATTTCTCAGTTGGTAAACC
AATCTTCTATCCATTCATTTTCGGATTCTCAATTTGGTCTTATTTTCGATTGGTGAAAACCGTCAAGCTTCAAGAAAA
CATATGGTTGTTGCCTTGAAAGAATTGAGTATTAGAGGTGATTTTGAAGTACTGTTGAGTATTTAATCAATTTATAGA
AACTCCAGATTTGAGGATAATACCATTAACCTGGTTGGTTGGATGAATTAATCACCAAAAGTTGACTGCTGAAAGAC
CAGATCCAATAGTTGCTGTTGTTGGAGCTGTAACCAAGCACACATCCAGGCTGAGGAAGAGAAAAGGAATACATC
40 CAATCTTTGAAAAAGGTCAAGTTCTCAGAGAACTTTGTAAGATGTTGTAAGTATTTTCCAGTTGAGTTTATTTATGAAGTGA
AAGATACAAGTTCACTGCTACTAAATCTTCAGAAGATAAATATACTTTGTTCTTAAATGGTTCTCGTTGTTGGTG
CACGTTCAATGTCGATGGTGGTTTATTTGTGTCATTAGATGGGAAATCACATTCTGTCTATTGGAAGGAAGAGGCATCT
GCCACTAGATTCAAGTTGATGGCAAACTGTTTATTAAGTTGAAAATGATCCACACAATTAAGAACTCCATCTCC
AGGTAAATTTGGTCAAGTATTTGGTTGACAGTGGTGAAGTCTGCTGCTGCTCAACCATACGCTGAAGTCAAGTTATGA
45 AAATGTGTATGCCTTTGATTGCTCAAGAAAAAGGGTAGTGAGTTGATTAACAACACCGGGTTCCACAGTTAATGCTGGT
GATATCTTGGCCATTTTGGCATTGGACGATCCATCTAAGGTCAAACATGCTAAACCATTTGAAGGTACTTTACCATCTAT
GGGTGAGCCAAATGTTACAGGTACTAAACCAGCACATAAATCAATCATTGTGCTGGTATTTTGAAGAACATTTTGGCTG
GTTATGATAATCAAGTGATTTTGAATTTCTACTTTAAAGAGTCTTGGTGAAGTTTGAAGACAATGAATTGCCATACTCT
GAATGGCAACAACAAATTCAGCTTTACACTCCAGATTGCCACCTAAATTTGGATGACGGATTGACTGATTGGTTGAAAG
50 AACTCAAAGTAGAGGTGCTGAATTCCTGCTCGTCAAATTTTAAACTCATCACCAAATCAATTGCTGAAAATGGTAATG
ATATGTTAGAAGATGTTGTTGCACCATTTGTTTCTATTGCCACAAGTTACCAGAATGGTTTGGTTGAACACGAATACGAT
TACTTTGCATCTTTGATTAAACGAATATTATGACGTTGAAAGTTTGGTTTTCAGGTGAAAATGTTAGAGAAGATAATGTTAT
CTTGAAATTAAGAGATGAAAACAAATCTGATTTGAAAAAAGTTATTGGTATTGGTTGTTCTCATTACGTTAGTGCCA
AGAACAATTTGATTTTAGCTATTTTGGACATTTATGAACCATTTGTTGAATCCAATCGTCAGTTGCTGCCTCTATCAGA
55 GAAGCTTTAAAGAAGTTGTTCAATAGACCTCGTGCTTGTGCCAAAGTTGCATTAAAGGCAAGAGAAATTTAATTCAATG
TTCTTTACCTTCCATCAAGGAAGATCCGATCAATTTGGAACATATTTTGAAGTCACTGTTGTTCAAACCTCTTATGGTG
AAATTTTGTCTAAACATAGAGAACCAATTTGGAATTTATTCGTGAGGTTGTTGATTCCAAACATATTGTTTATGTTG
TTGGCACAATTTCTAATCAATCCAGACCCATGGGTTGCCATTGCTGCCGCTGAAGTTTATGTCAGACGTTTCATACCGTGC
TTATGATTTGGGTAAAATTAATATCATGTTAATGACAGACTTCTTATTTGTTGAATGGAATTCAGTTGGCTAATATGG
60 GAGCCGCTGGTGTAAACGATGCTCAACAGGCTGCTGCTGCCGCTGGGATGATTCGACATCTATGAAACATGCAGCTTCT
GTGTCTGATTTGACCTTTGTTGTTGATTTCTAAAACCGAGCATTCCACAAGAAGTGGTGTGTTAGCTCCAGCAAGACACTT
GGATGATGTTGATGAAACTCTTACAGCTGCATTGGAACAAATCCAACAGCCGATGCTATTTTCAATTAAAGCAAAGGGTG
AAACTCCAGAGTTATTAATGTTTTGAATATTGTCATTACCAGTATTGATGGTTACTCCGATGAAAATGAATACTTGAGC
AGAATTAATGAAATCTTGTGCGAATACAAAGAAGAGTTGATTTCTGCTGGTGTGCTCGTGTACATTTGTTTTGCTCA

TCAAATTGGTCAATATCCTAAATATTATACTTTTACTGGTCTGACTATGAAGAAAACAAGGTTATTAGACACATTGAAC
CAGCTTTGGCTTTCCAATTGGAATTGGGAAGATTAGCCAATTTTCGATATCAAACCAATTTTCACTAACCAACAGAAACATC
CATGTATATGATGCAATTGGGAAGAATGCTCCTTCTGATAAAAGATTTTACCAGAGGGATTATTAGAACCAGGTTGTTCT
TAAAGAAGACATTAGCATTAGTGAATATTTGATTGCTGAATCCAACAGATTAATGAATGATATTTGGATACTTTAGAAG
5 TTATTGACACTTCCATTTCTGATTTAAACCATATTTTCACTTAACCTTTTCCAATGCTTTCAATGTTCAAGCTTCAGATGTT
GAGGCTGCCTTTGGATCATTCTTAGAAAGATTTGGTAGAAGATTATGGAGATTAAGAGTTACTGGTGCTGAAATTAGAAT
TGTCTGTACTGATCCTCAAGGTAATTCGTTCCCATTCGCTGCTATCATTAAATAATGTTTCTGGTTATGTTGTCAAATCAG
AATTGTATTTGGAAGTGAAAAATCCTAAAGGTGAATGGGTTTTCAAATCCATTGGTCATCCTGGTTCCATGCATTTGAGA
CCTATCTCAACTCCATATTCGTTTAAAGAATCTTTACAACCAAAACGTTACAAGGCTACAATATGGGTACCACTTATGT
10 GTATGACTTCCCAGAATTGTTTCGTCAAGCAACAATTTACAATGGAAAAAATATGGCAAAAAAGTACCAAAAGATGTTT
TCGTGTCTTTAGAATTGATCACTGATGAAACTGATTCCTTAATAGCTGTTGAAAGAGATCCGGGTGCTAACAAAATTGGA
ATGTTGGATTCAAAGTCACTGCTAAAACCTCCTGAATACCCTCATGGTCCGTCATTAATTATGTTGCCAATGATATCAC
CCACAAGATTGGTTCTTTTGGTCCAGAAGAAGATTAATTTCAACCAAGTGTACTGAATTTGGCCAGAAAATTAGGTATTC
CAAGAATTTACCTTTCTGCAAAATTCAGGTGCTAGAATTGGTGTGCTGAGGAATTGATTCCATTATACCAAGTTGCCCTGG
15 AATGAAGAAGGGTCTCCTGACAAAGGATTAGATCTTGTACTTGAGTACTGCTGCTAAAGAGTCTTTAGAAAAAGATGG
TAAAAGTGACAGTGTGTTACTGAACGTATTGTTGAAAAAGGTGAAGAGCGTCATGTCATTAAAGCTATTATTGGTGCCG
AAGATGGCTTAGGGGTTGAATGTCAAAGGATCAGGTTAATTGCTGGTGCCACATCAAGAGCTTACAAGGATATATTT
ACCATCACTTTGGTAACCTGTAGATCTGTTGGTATTGGTGCTTATTGGTTAGATTGGGTCAAAGAGCTTCAAAATCGA
TGGTCAACCTATTATTTTAACTGGTGCTCCTGCTATCAATAAATTGTTGGGTAGAGAAGTGTATTCTTCCAATCTTCAAT
20 TGGGTGGTACTCAAATCATGTACAATAATGGTGTCTCTCATTGACAGCTAATGATGATTGGGTGGGGTTGAAAAAATT
ATGGAATGGTTATCATATGTTCCAGCTAAACGTGGTTTACCAGTGCCAATTTTGAATCAGAAGATTCTTGGGACAGAGA
TGTTGATTACTACCCACCAAAAACAAGAAGCTTTTGATGTTAGATGGATGATCCAAGGTAGAGAAGTTGATGGTGAATATG
AATCTGGGTATTATTTGATAAAGATTCAATCCAGAAACATTATCTGGTTGGGCTAAAGGTGTTGTTGTTGGTAGAGCACGT
TTGGGTGGTATTCCAATTGGTGTATTGGTGTGCGAAACCAGAACAGTGGAAAACCTGATTCCTGCTGATCCAGCAAATCC
25 AGACTCTACAGAAAGTTGATTCAAGAAGCAGGTCAAGTGTGGTATCCTAACTCTGCTTTTAAAGACAGCACAAAGCTATAA
ATGATTTCAACAATGGTGAACAATTGCCATTAATGATTTTAGCAAATTGGAGAGGTTTCTCTGGTGGTCAAAGAGATATG
TACAATGAAGTCTTGAAATATGGTTCAATTTATGTTGATGCTTTAGTTGACTTCAAGCAACCTATCTTCACTTACATTCC
ACCAAATGGAGAATTGAGAGGTGGCTCTTGGGTTGTTGTTGATCCAACCATCAACTCAGATATGATGGAAATGTATGCCG
ATGTCGATTTCGAGAGCTGGTGTGTTTGGAAACCAGAAGGTATGGTTGGTATCAAATACAGACGTGATAAATTATTAGCAACT
30 ATGGAAAGATTAGATCCAACCTTATGGTGAATGAAAGCTAAGTTAAATGACTCGTCAATATCTCCAGAAGAACACTCGAA
AATAAGCGCCAAATTTGTTGCACGTGAAAGGCTTTATTACCAATTTATGCTCAAATTTCCGTTCAATTTGCTGACTTGC
ACGATAGATCAGSTCGTATGTTGGCCAGGGAGTTATTAGAAAGGAAATCAAATGGACTGATGCTAGACGTTTCTTCTTC
TGGAGATTGAGAAGAAGATTGAACGAGGAATATGTTTTGAGATTGATTAGTGAACAAATTAAGATTCTAGCAAATTGGA
AAGACTTGCCAGATTGAAGAGTTGGATGCCAATGTTGAATACGATGATGACCAAGCTGTCACTAAGTGAAGAGA
35 ACCATGCCAAATTTGCAAAAGAGAGTTAATGAATTGAAACAAGAAGTTTCAAGAACCAAGATTATGAGATTATTAAGAGAG
GATCCAAATAGTGCAATTTCTGCAATGAAAGACTATGTTGAAAGATTGTCAAAGAGATAAAGAGAAATTCCTCAAGGC
ATTGAAGtagAAGTGGTTTCCATTAATTCAACTTTTTAATGACATTGAAAGTAGTAGTGTGTTTTTTAGATTTAA
GTATATTATATTATGAATAAATTATAGAAAGTAATTATAGTTTTGACGGTTAATTGACGAGAGTGGGAAATTGGCTTTT
40 TTGTTGCTCGTGTGATGAAACAGTGATTGACACAAAAAATAGACAATGAAAC

FIGURE 5

5 MRCKLSLIKNTNSLVHRSRFLITKPQLYIPHRHYIPFKNIFKSLMSDQSPSPSPSDLSYTTLHENLP SHFLGGNSVLN
AEP SKVRDFVRAHQGHTVISKILIANNGIAAVKEIRSVRKWAYETFGDEKAIQFTVMATPEDLEANA EYIRMA DQFIEVP
GGTNNNNYANVDLIVEIAESTNAHAVWAGWGHASENPLLPEKLAASPKKII FIGPPGSAMRSLGDKISSTIVAQHAQVPC
IPWSGTGVDVKIDPQTNLVS VADDIYAKGCCTSPEDGLEKAKKIGFPVMIKASEGGGGKGIRKVDDEKNFITLYNQAAN
EIPGSPIFIMKLAGDARHLEVQLLADQYGTNISLFG RDCSVQRRHQKIIIEEAPVTIARKET FHEMENA AAVRLGKLVGYVS
10 AGTVEYLYSHAEDKFYFLELNPRLQVEHPTTEMVTGVNLPAAQLQIAMGIPMHRIRDIRTLYGADPHTTTTDIDFEFKSET
SLVSQRRTPKGHCTACRITSEDPGEGFKPSGGS LHELNFRSSSNVWGYFSVGNQSSIH SFSDSQFGHIFAFGENRQASR
KHMVVALKELSIRGDFRTTVEYLIKLETPDFEDNTITTGWLDELITKKLTAERPDP I VAVVCGAVTKAHIQAE EEEKKEY
IQSLEKGQVPHRNLLKTIFPVEFIYEGERYKFTATKSS EDKYTLFNLGSR CVVGARSLS DGGLLCALDGKSHSVYWK EEA
SATRLSVDGKTCLEVEN DPTQLRTPSPGKLVKYLVDSEHVDAGOPYAEVEVMKMCPLIAQENG VVQLIKQPGSTVNA
15 GDILAILALDDPSKVKHAKPFEGTLP SMGEPNVTGTKPAHKFNHCAGILKNILAGYDNQVILNSTLKS LGEVLKDNELPY
SEWQQQISALHSRLPPKLD DGLTALVERTQSRGAEFFPARQILKLITKSIAENGND MLEDV VAPLVSIATSYQNG LVEHEY
DYFASLINEYYDVESL FSGENVREDNVILKLRDENKSDLKKVIGIGLSHSRVS AKNNLILAILDIYE PLLQSNSSVAASI
REALKNLFIRPRACAKVALKAREILIQCSLPSIKERSDQLEHILRSSV VQTSYGEI FAKHREP NLEI IREVVDSKHIVFD
VLAQFLINPD PWVAIAAAEVYVRRSYRAYDLGKIEYHVNDRLPIVEWKFKLANMGAAGV NDAQAAAAGDDSTSMKHAA
SVSDLTFFVDSKTEHSTR TGVLAPARHLDDVDET LTAALEQFQPADAI SFKAKGETPELLNVLNIVITSIDGYS DENEYL
20 SRINEILCEYKEELISAGVRRVT FVFAHQIGQYPKYYTFTGPDYEENKVIRHIEPALAFQELGRLANFDIKPIFTNNRN
IHVYDAIGKNAPSDKRFFTRGIIRTGV LKEDISISEYLIAESNR LNMNDILDTLEVIDT SNSDLNHIFINFSNAFN VQASD
VEAAFGSFLERFGRRLWRLRV TGAEIRIVCTDPQGT SFPLRAIINNVS GYVVKSELYLEVKNPKGEWVFKSIGHPGSMHL
RPISTPYPVKESLQPKRYKAHNMGTTYVYDFPELFRQATISQWK KYGKKVPKDV FVSLELITDETDSLIAVERDPGANKI
GMVGFKVTAKTPEYPHGRQLIIVANDITHKIGSFGPEEDNYFNKCTELARKLGI PRIYLSANS GARIGVAEELIPLYQVA
25 WNEEGSPDKGFRYLYLSTA AKESLEKDGKSDSVTERIVEKGEERHVIKAIIGAEDGLGVECLKGSGLIAGATSRAYKDI
FTITLVTCRSVGIGAYLVRLGQRAIQIDGQPIILTGAPAINKLLGREVYSSNLQLGGTQIMYNNGVSHLTANDDLAGVEK
IMEWLSYVPAKRGLPVPILESEDSWDRD VDYPPKQEA FDVWRWMIQGREVDGEYESGLFDKDSFQETLSGWAKGVVVGRA
RLGGIPIGVIGVETRTVENLIPADPANPDSTESLIQEAGQVWYPNSAFKTAQA INDFNNGEQLPLMILANWRGFSGGQRD
MYNEVLKYGSFIVDALVDFKQPIFTYIPNGELRGGSWVVDPTINSDMMEMYADVDSRAGVLEPEGMVG IKYRRDKLLA
30 TMERLDPTYGEMKAKLNDSSLSPEEHSKISAKLFAREKALLPIYAQISVQFADLHDSGRMLAKGVIRKEIKWTDARRFF
FWRLRRRLNEEYVLR LISEQIKDSSKLERVARLKSWMPTVEYDDDQAVSNWIEENHAKLQKRVNELKQEVSR TKIMRLK
EDPNSAISAMKDYVERLSKEDKEKFLKALK

35

7/8

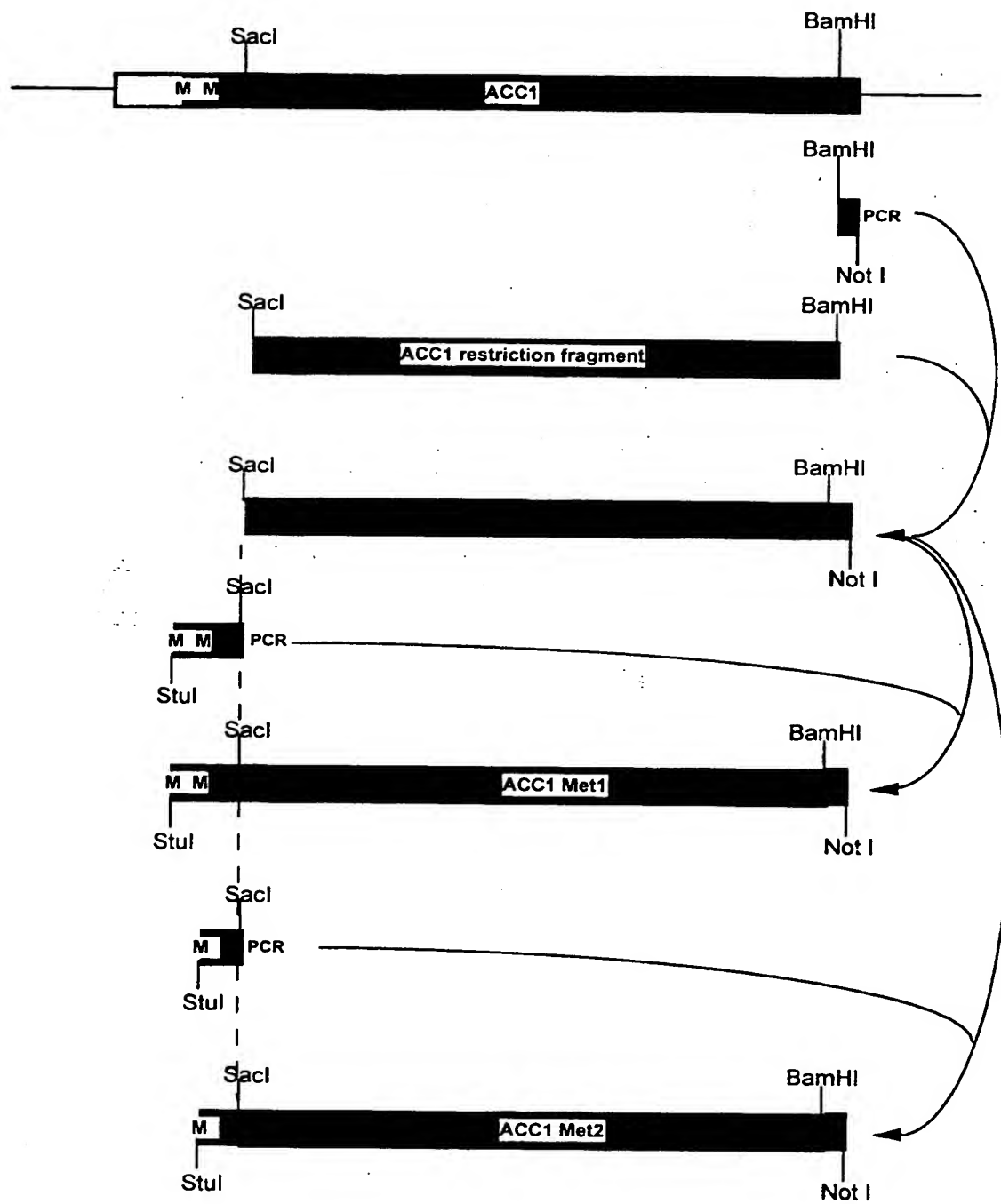
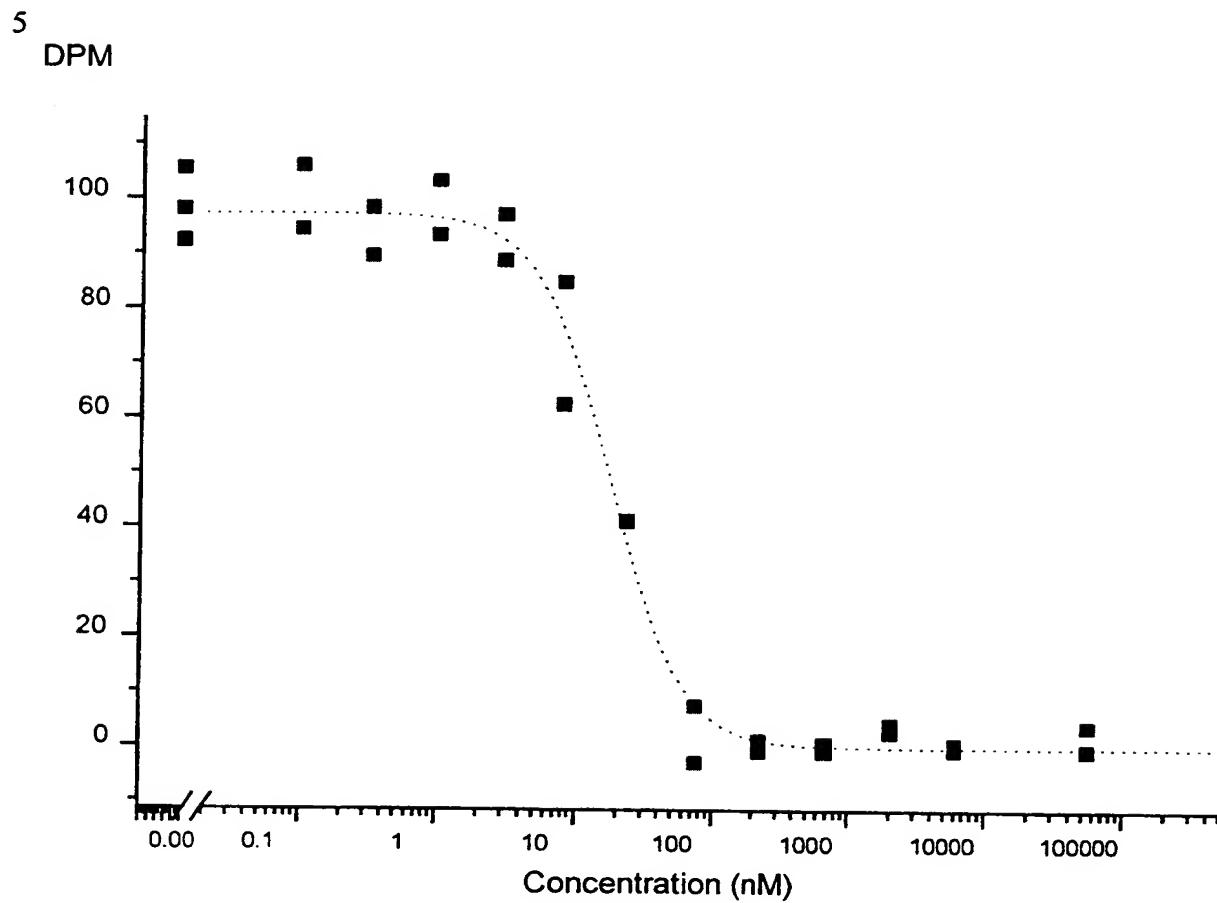
FIGURE 6

FIGURE 7

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT:

- (A) NAME: Zeneca Ltd
(B) STREET: 15 Stanhope Gate
(C) CITY: London
(D) STATE: Greater London
(E) COUNTRY: England
(F) POSTAL CODE (ZIP): W1Y 6LN
(G) TELEPHONE: 0171 304 5000
(H) TELEFAX: 0171 304 5151
(I) TELEX: 0171 834 2042

10

15

(ii) TITLE OF INVENTION: PROCESS

(iii) NUMBER OF SEQUENCES: 3

20

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

25

(vi) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: GB 9726897.3
(B) FILING DATE: 20-DEC-1997

30

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 523 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: other nucleic acid

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCACGCTTGA CGGTTTTTCAC CAAATGCGAA AATATGACCA AATTGAGAAT CCGAAAATGA

60

ATGGATAGAA GATTGGTTAC CAACTGAGAA ATAACCCAC ACATTAGAAG AAGAACGGAA 120
 ATTCAATTCA TGTAAGAAC CACCACTTGG TTTAAACCT TCACCAGGAT CTCAGAAGT 180
 AATACGACAA GCAGTACAAT GTCCCTTGG TGTGGTCTT CTTTACTAA CCAATGAAGT 240
 TTCTGACTTG AATTCAAAAT CAATATCAGT AGTGGTATGA GGATCGGCAC CGCACAAAGT 300
 5 TCTGATATCT CTGATTCTAT GCATTGGTAT ACCCATAGCA ATTTGTAATT GAGCAGCTGG 360
 TAAATTAACA CCTGTCACCA TTTCAGTGGT TGGATGTTCA ACTTGCAATC TTGGGTTCAA 420
 TTCCAAAAG TAGAATTTAT CTCAGCGTG GGGAGTAAAG GACTCAACA GTACCAGGGG 480
 GTTACATAAC CAACTTATTT TACCCAATCT GACTGGTGGA TTT 523

10 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8054 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

25 AATATATTGC TTCCTTTTGA TAGGAAGTAA CTCCGAGTGT TTGAATTTGA TATATGTTAT 60
 TCATATACGT TCAATGGCTC TCTTCTATGC TTTGTATATA CTTTCTTTTG AATAGATACT 120
 CATGTAAAGA GATTTGAAAC CATATTCTAA CCAACAAAAA TATTGTACGG TATAGGTTAG 180
 AAAAAAACT CCGTAAGGTC CGCTTACACG GTTAAATTGA AAACACGTTA AAAATATATT 240
 TGGGTAATGG ACTAAGCTAT ATACAGTACT CAACAAAAAT GAAATCAAAC ACAATGTTCT 300
 30 TTGGGAAATT CATTTTCATGC AACTAGGGTG ATTCTCTTTC TACTATCCAA CAACGATAAC 360
 CCTGCTTTTG AAAAATCTTT TCTAAATTCA AATTGATATA ATTCTATTTT ATATATTACT 420
 TTCTTTTCC CATATAACCC CATTTTTTTT TTGGAATCAT ATTTGTTTTT GATTTTTGCT 480
 TTCCCTTTCA GTCTGAGGAA CATACTAATT ACGAACAACA ATTATACATC CAATCTTCAT 540
 CTAACGAATT GATTATTTAC ATTTATTAAA CCCTTGGATA CAACTGATT ACACTTTTTA 600
 35 GTTAGTTTGT TCAATTATAA GGGTATTATA CAACAAAGAT ATCATTTAAA GTTAAATCTC 660
 AATCTGGAAT AATAAAAGTA TTCAACACTT TTGCTTACAA TAGGTATGTT CAAAATCAAT 720
 TGAAGCCATC GAGATAAGAA ATTAAGCAAA AACGTTTACA ATTGTTGTGT GTGTGTTGCA 780
 GTGTTTGAAG AAGCTCGAGT GATTGCTTTT CTTCGGCATC AGCTGTGTTG GGAACATCTT 840
 GTCGTTAAAG TTTCCGAGTA ATATTAGAGT AATGGAACGA AAAAAACAAA ATAAAGTTCT 900
 40 GGAACCACAA AGATTGAAA AATTGGGTAG AAACAAAAA AAGACAAAGC AGGAACCCAA 960
 CAATAAATGA ATAAACACTC AAAAATACT CACAACAACA ACACTTATTT TCACTTGCTT 1020
 TATTTCTTCG ATTTTTTATG AGATGCAAAAT TATCTCTAAT AAAGAATACT AACTCACTTG 1080
 TACATAGATC GCGTTTCCTA ATTACAAAAC CAACTATA TATACCTCAT CGTCATTATA 1140
 TCCCATTCOA GAACATATTC AAGTCATTGT TAATGTCAGA TCAATCTCCA TCTCCTAGTC 1200
 45 CTAGCGATTG CCTTAGCTAC ACTACATTAC ATGAAAATTT GCCATCTCAT TTCTGGGTG 1260
 GAAATTCAGT TTTGAATGCT GAACCTTCTA AAGTCAGAGA CTTTGTGAGA GCTCATCAAG 1320
 GTCATACAGT TATTTGAAA ATTTTAATTG CCAACAATGG TATAGCTGCA GTTAAAGAAA 1380

	TCAGATCAGT TAGAAAATGG GCTTATGAAA CATTGTTGGA CGAAAAAGCC ATACAGTTTA	1440
	CCGTTATGGC CACTCCAGAA GATTGGAAG CTAATGCCGA ATATATTAGA ATGGCCGACC	1500
	AATTCATTGA AGTCCCTGGT GGCACCAATA ACAATAACTA TGCTAATGTT GATCTCATTG	1560
	TAGAGATAGC AGAAAGTACA AATGCTCATG CCGTTTGGGC TGGGTGGGG CATGCTTCAG	1620
5	AGAATCCTTT GTTACCAGAA AAATTAGCTG CATCTCCAA AAAAATTATT TTTATTGGTC	1680
	CTCCTGGTTC AGCTATGAGA TCTTTAGGTG ACAAGATTTC ATCTACTATA GTTGCTCAAC	1740
	ATGCTCAAGT ACCATGTATT CCATGGTCCG GTAGTGGTGT TGATGAAGTG AAAATAGACC	1800
	CACAACTAA TTTGGTTTCT GTTGCTGATG ATATTTATGC CAAAGGGTGC TGTACTAGTC	1860
	CAGAAGATGG TTTAGAAAAA GCCAAAAAAA TTGGGTTCCC AGTTATGATT AAAGCCTCTG	1920
10	AAGGTGGTGG TGGTAAAGGT ATTAGAAAAG TTGATGATGA GAAAACTTC ATTACCTTAT	1980
	ACAACCAAGC AGCTAATGAA ATACCAGGTT CTCCTATCTT TATTATGAAG TTAGCAGGTG	2040
	ATGCCAGACA TTTAGAAGTT CAATTACTAG CAGATCAATA CGGTACTAAC ATTTCCCTTT	2100
	TTGGAAGAGA TTGTTCCGTA CAAAGAAGAC ACCAAAAGAT TATTGAAGAA GCACCAGTCA	2160
	CCATTGCCAG AAAGGAACT TTCCACGAAA TGGAAAATGC AGCAGTCAGA TTGGGTAAAT	2220
15	TAGTTGGTTA TGTATCCGCT GGTACTGTTG AGTATCTTTA CTCCCACGCT GAAGATAAAT	2280
	TCTACTTTTT GGAATTGAAC CCAAGATTGC AAGTTGAACA TCCAACCACT GAAATGGTGA	2340
	CAGGTGTTAA TTTACCAGCT GCTCAATTAC AAATTGCTAT GGGTATACCA ATGCATAGAA	2400
	TCAGAGATAT CAGAACTTTG TACGGTGCCG ATCCTCATAC CACTACTGAT ATTGATTTTG	2460
	AATTCAGTC AGAACTTCA TTGGTTAGTC AAAGAAGACC AACACCAAAG GGACATTGTA	2520
20	CTGCTTGTCG TATTACTTCT GAAGATCCTG GTGAAGGTTT TAAACCAAGT GGTGGTTCTT	2580
	TACATGAATT GAATTTCCGT TCTTCTTCTA ATGTGTGGGG TTATTTCTCA GTTGGTAACC	2640
	AATCTTCTAT CCATTCATTT TCGGATTCTC AATTTGGTCA TATTTTCGCA TTTGGTAAA	2700
	ACCGTCAAGC TTCAAGAAAA CATATGGTTG TTGCCTTGAA AGAATTGAGT ATTAGAGGTG	2760
	ATTTTAGAAC TACTGTTGAG TATTTAATCA AATTGTTAGA AACTCCAGAT TTCGAGGATA	2820
25	ATACCATTAC AACTGGTTGG TTGGATGAAT TAATCACCBA AAAGTTGACT GCTGAAAGAC	2880
	CAGATCCAAT AGTTGCTGTT GTTGTGGAG CTGTAACCAA AGCACACATC CAGGCTGAGG	2940
	AAGAGAAAAA GGAATACATC CAATCTTTGG AAAAAGGTCA AGTTCCTCAC AGAAACTTAT	3000
	TGAAAATAT TTTCCAGTT GAGTTTATTT ATGAAGGTGA AAGATACAAG TTCACTGCTA	3060
	CTAAATCTTC AGAAGATAAA TATACTTTGT TCCTTAATGG TTCTCGTTGT GTTGTGGTG	3120
30	CACGTTTATT GTCCGATGGT GGTTTATTGT GTGCATTAGA TGGGAAATCA CATTCTGTCT	3180
	ATTGGAAGGA AGAGGCATCT GCCACTAGAT TATCAGTTGA TGGCAAACT TGTTTATTAG	3240
	AAGTTGAAAA TGATCCAACA CAATTAAGAA CTCCATCTCC AGGTAAATTG GTCAAGTATT	3300
	TGGTTGACAG TGGTGAACAT GTTGATGCTG GTCAACCATA CGCTGAAGTC GAAGTTATGA	3360
	AAATGTGTAT GCCTTTGATT GCTCAAGAAA ATGGGGTAGT GCAGTTGATT AAACAACCGG	3420
35	GTTCCACAGT TAATGCTGGT GATATCTTGG CCATTTTGGC ATTGGACGAT CCATCTAAGG	3480
	TCAAACATGC TAAACCATTG GAAGGTACTT TACCATCTAT GGGTGAGCCA AATGTTACAG	3540
	GTAATAAACC AGCACATAAA TTCAATCATT GTGCTGGTAT TTTGAAAAAC ATTTTGGCTG	3600
	GTTATGATAA TCAAGTGATT TTGAATTCTA CTTTAAAGAG TCTTGGTGAA GTTTTGAAAG	3660
	ACAATGAATT GCCATACTCT GAATGGCAAC AACAAATTC AGCTTTACAC TCCAGATTGC	3720
40	CACCTAAATT GGATGACGGA TTGACTGCAT TGGTTGAAAG AACTCAAAGT AGAGGTGCTG	3780
	AATCCCTGTC TCGTCAAATT TTAATACTCA TCACCAAATC AATTGCTGAA AATGGTAATG	3840
	ATATGTTAGA AGATGTTGTT GCACCATTGG TTTCTATTGC CACAAGTTAC CAGAAATGGT	3900
	TGGTTGAACA CGAATACGAT TACTTTGCAT CTTTGATTAA CGAATATTAT GACGTTGAAA	3960
	GTTTGTTTTC AGGTGAAAAT GTTAGAGAAG ATAATGTTAT CTTGAAATTA AGAGATGAAA	4020
45	ACAAATCTGA TTTGAAAAAA GTTATTGGTA TTGGTTTGTG TCATTACAGT GTTAGTGCCA	4080
	AGAACAATTT GATTTTAGCT ATTTTGGACA TTTATGAACC ATTGTTGCAA TCCAACCTCGT	4140
	CAGTTGCTGC CTCTATCAGA GAAGCTTTAA AGAACTTGTT CATTAGACCT CGTGCTTGTG	4200

	CCAAAGTTGC	ATTAAAGGCA	AGAGAAATTT	TAATTCAATG	TTCTTTACCT	TCCATCAAGG	4260
	AAAGATCCGA	TCAATTGGAA	CATATTTTGA	GGTCATCTGT	TGTTCAAACC	TCTTATGGTG	4320
	AAATTTTTGC	TAAACATAGA	GAACCAAATT	TGGAAATTAT	TCGTGAGGTT	GTTGATTCCA	4380
	AACATATTGT	TTTTGATGTG	TTGGCACAAT	TCTTAATCAA	TCCAGACCCA	TGGGTTGCCA	4440
5	TTGCTGCCGC	TGAAGTTTAT	GTCAGACGTT	CATACCGTGC	TTATGATTTG	GGTAAAATTG	4500
	AATATCATGT	TAATGACAGA	CTTCCTATTG	TTGAATGGAA	ATTCAAGTTG	GCTAATATGG	4560
	GAGCCGCTGG	TGTAAACGAT	GCTCAACAGG	CTGCTGCTGC	CGGTGGCGAT	GATTCGACAT	4620
	CTATGAAACA	TGCAGCTTCT	GTGTCTGATT	TGACCTTTGT	TGTTGATTCT	AAAACCGAGC	4680
	ATTCCACAAG	AACTGGTGTT	TTAGCTCCAG	CAAGACACTT	GGATGATGTT	GATGAAACTC	4740
10	TTACAGCTGC	ATTGGAACAA	TTCCAACCAG	CCGATGCTAT	TTCAATTTAAA	GCAAAGGGTG	4800
	AAACTCCAGA	GTTATTAAAT	GTTTTGAATA	TTGTCAATTAC	CAGTATTGAT	GGTTACTCCG	4860
	ATGAAAATGA	ATACTTGAGC	AGAATTAATG	AAATCTTGTTG	CGAATACAAA	GAAGAGTTGA	4920
	TTTCTGCTGG	TGTTCTGCTG	GTTACATTTG	TTTTTGCTCA	TCAAATTGGT	CAATATCCTA	4980
	AATATTATAC	TTTTACTGGT	CCTGACTATG	AAGAAAACAA	GGTTATTAGA	CACATTGAAC	5040
15	CAGCTTTGGC	TTTCCAATTG	GAATTGGGAA	GATTAGCCAA	TTTCGATATC	AAACCAATTT	5100
	TCACTAACAA	CAGAAACATC	CATGTATATG	ATGCAATTGG	GAAGAATGCT	CCTTCTGATA	5160
	AAAGATTTTT	CACCAGAGGG	ATTATTAGAA	CCGGTGTTCT	TAAAGAAGAC	ATTAGCATT	5220
	GTGAATATTT	GATTGCTGAA	TCCAACAGAT	TAATGAATGA	TATTTTGGAT	ACTTTAGAAG	5280
	TTATTGACAC	TTCTAATTCT	GATTTAAACC	ATATTTTCAT	TAACTTTTCC	AATGCTTTCA	5340
20	ATGTTCAAGC	TTCAGATGTT	GAGGCTGCCT	TTGGATCATT	CTTAGAAAGA	TTTGGTAGAA	5400
	GATTATGGAG	ATTAAGAGTT	ACTGGTGCTG	AAATTAGAAT	TGTCTGTACT	GATCCTCAAG	5460
	GTACTTCGTT	CCCATTGCGT	GCTATCATT	ATAATGTTTC	TGGTTATGTT	GTCAAATCAG	5520
	AATTGTATTT	GGAAGTGAAA	AATCCTAAAG	GTGAATGGGT	TTTCAAATCC	ATTGGTCATC	5580
	CTGGTTCCAT	GCATTTGAGA	CCTATCTCAA	CTCCATATCC	AGTTAAAGAA	TCTTTACAAC	5640
25	CAAAACGTTA	CAAGGCTCAC	AATATGGGTA	CCACTTATGT	GTATGACTTC	CCAGAATTGT	5700
	TTCGTCAAGC	AACAATTTCA	CAATGGAAAA	AATATGGCAA	AAAAGTACCA	AAAGATGTTT	5760
	TCGTGTCTTT	AGAATTGATC	ACTGATGAAA	CTGATTCCTT	AATAGCTGTT	GAAAGAGATC	5820
	CGGGTGCTAA	CAAAATTGGA	ATGGTTGGAT	TCAAAGTCAC	TGCTAAAACT	CCTGAATACC	5880
	CTCATGGTCG	TCAATTAATT	ATTGTTGCCA	ATGATATCAC	CCACAAGATT	GGTTCTTTTG	5940
30	GTCCAGAAGA	AGATAATTAT	TTCAACAAGT	GTAATGAATT	GGCCAGAAAA	TTAGGTATTC	6000
	CAAGAATTTA	CCTTTCTGCA	AATTCAGGTG	CTAGAATTGG	TGTTGCTGAG	GAATTGATTC	6060
	CATTATACCA	AGTTGCCTGG	AATGAAGAAG	GGTCTCCTGA	CAAAGGATTC	AGATACTTGT	6120
	ACTTGAGTAC	TGCTGCTAAA	GAGTCTTTAG	AAAAAGATGG	TAAAAGTGAC	AGTGTGTGTA	6180
	CTGAACGTAT	TGTTGAAAAA	GGTGAAGAGC	GTCATGTCAT	TAAAGCTATT	ATTGGTGCCG	6240
35	AAGATGGCTT	AGGGGTTGAA	TGTCTTAAAG	GATCAGGTTT	AATTGCTGGT	GCCACATCAA	6300
	GAGCTTACAA	GGATATATTT	ACCATCACTT	TGGTAACTTG	TAGATCTGTT	GGTATTGGTG	6360
	CTTATTTGGT	TAGATTGGGT	CAAAGAGCCA	TTCAAATCGA	TGGTCAACCT	ATTATTTTAA	6420
	CTGGTGCTCC	TGCTATCAAT	AAATTGTTGG	GTAGAGAAGT	GTATTCTTCC	AATCTTCAAT	6480
	TGGGTGGTAC	TCAAATCATG	TACAATAATG	GTGTTTCTCA	TTTGACAGCT	AATGATGATT	6540
40	TGGCTGGGGT	TGAAAAAATT	ATGGAATGGT	TATCATATGT	TCCAGCTAAA	CGTGGTTTAC	6600
	CAGTGCCAA	TTTGAATCA	GAAGATTCTT	GGGACAGAGA	TGTTGATTAC	TACCCACCAA	6660
	AACAAGAAGC	TTTTGATGTT	AGATGGATGA	TCCAAGGTAG	AGAAGTTGAT	GGTGAATATG	6720
	AATCTGGGTT	ATTTGATAAA	GATTCATTCC	AAGAAACATT	ATCTGGTTGG	GCTAAAGGTG	6780
	TTGTTGTTGG	TAGAGCACGT	TTGGGTGGTA	TTCCAATTGG	TGTTATTGGT	GTCGAAACCA	6840
45	GAACAGTGG	AAACTTGATT	CCTGCTGATC	CAGCAAATCC	AGACTCTACA	GAAAGTTTGA	6900
	TTCAAGAAGC	AGGTCAAGTG	TGGTATCCTA	ACTCTGCTTT	TAAGACAGCA	CAAGCTATAA	6960
	ATGATTTCAA	CAATGGTGAA	CAATTGCCAT	TAATGATTTT	AGCAAATTGG	AGAGGTTTCT	7020

-5-

CTGGTGGTCA AAGAGATATG TACAATGAAG TCTTGAAATA TGGTTCATTT ATTGTTGATG 7080
 CTTTAGTTGA CTTCAAGCAA CCTATCTTCA CTTACATTCC ACCAAATGGA GAATTGAGAG 7140
 GTGGCTCTTG GGTGTGTGTT GATCCAACCA TCAACTCAGA TATGATGGAA ATGTATGCCG 7200
 ATGTCGATTC GAGAGCTGGT GTTTTGGAAC CAGAAGGTAT GGTGGGTATC AAATACAGAC 7260
 5 GTGATAAATT ATTAGCAACT ATGGAAAGAT TAGATCCAAC TTATGGTGAA ATGAAAGCTA 7320
 AGTTAAATGA CTCGTCATTA TCTCCAGAAG AACACTCGAA AATAAGCGCC AAATTGTTTG 7380
 CACGTGAAAA GGCTTTATTA CCAATTTATG CTCAAATTTT CGTTCAATTT GCTGACTTGC 7440
 ACGATAGATC AGGTCGTATG TTGGCCAAGG GAGTTATTAG AAAGGAAATC AAATGGACTG 7500
 ATGCTAGACG TTTCTTCTTC TGGAGATTGA GAAGAAGATT GAACGAGGAA TATGTTTTGA 7560
 10 GATTGATTAG TGAACAAATT AAAGATTCTA GCAAATTGGA AAGAGTTGCC AGATTGAAGA 7620
 GTTGGATGCC AACTGTTGAA TACGATGATG ACCAAGCTGT CAGTAACTGG ATTGAAGAGA 7680
 ACCATGCCAA ATTGCAAAAG AGAGTTAATG AATTGAAACA AGAAGTTTCA AGAACCAAGA 7740
 TTATGAGATT ATTAAGAGAG GATCCAAATA GTGCAATTTT TGCAATGAAA GACTATGTTG 7800
 AAAGATTGTC AAAAGAAGAT AAAGAGAAAT TCCTCAAGGC ATTGAAGTAG AAGTGGTTTC 7860
 15 CATTAATTCA ACTTTTAAAT GACATTGAAA GTAGTAGTAG TTGTTGTTTT TTAGATTAA 7920
 GTATATTATA TTATGTAATA AATTATAGAA AGTAATTATA GTTTTGACGG TTAATTGACG 7980
 AGAGTGGGAA ATTGGCTTTT TTGTTGCTCG TGTGATGAAA CAGTGATTGA CACAAAAAAA 8040
 TAGACAATGA AAC 8054

20 (2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2270 amino acids
 (B) TYPE: amino acid
 25 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

35 Met Arg Cys Lys Leu Ser Leu Ile Lys Asn Thr Asn Ser Leu Val His
 1 5 10 15
 Arg Ser Arg Phe Leu Ile Thr Lys Pro Gln Leu Tyr Ile Pro His Arg
 20 25 30
 His Tyr Ile Pro Phe Lys Asn Ile Phe Lys Ser Leu Leu Met Ser Asp
 35 40 45
 40 Gln Ser Pro Ser Pro Ser Pro Ser Asp Ser Leu Ser Tyr Thr Thr Leu
 50 55 60
 His Glu Asn Leu Pro Ser His Phe Leu Gly Gly Asn Ser Val Leu Asn
 65 70 75 80
 45 Ala Glu Pro Ser Lys Val Arg Asp Phe Val Arg Ala His Gln Gly His
 85 90 95
 Thr Val Ile Ser Lys Ile Leu Ile Ala Asn Asn Gly Ile Ala Ala Val

	100		105		110
	Lys Glu Ile Arg Ser Val Arg	Lys Trp Ala Tyr Glu Thr Phe Gly Asp			
	115	120	125		
5	Glu Lys Ala Ile Gln Phe Thr Val Met Ala Thr Pro Glu Asp Leu Glu				
	130	135	140		
	Ala Asn Ala Glu Tyr Ile Arg Met Ala Asp Gln Phe Ile Glu Val Pro				
	145	150	155		160
	Gly Gly Thr Asn Asn Asn Asn Tyr Ala Asn Val Asp Leu Ile Val Glu				
	165	170	175		
10	Ile Ala Glu Ser Thr Asn Ala His Ala Val Trp Ala Gly Trp Gly His				
	180	185	190		
	Ala Ser Glu Asn Pro Leu Leu Pro Glu Lys Leu Ala Ala Ser Pro Lys				
	195	200	205		
15	Lys Ile Ile Phe Ile Gly Pro Pro Gly Ser Ala Met Arg Ser Leu Gly				
	210	215	220		
	Asp Lys Ile Ser Ser Thr Ile Val Ala Gln His Ala Gln Val Pro Cys				
	225	230	235		240
	Ile Pro Trp Ser Gly Thr Gly Val Asp Glu Val Lys Ile Asp Pro Gln				
	245	250	255		
20	Thr Asn Leu Val Ser Val Ala Asp Asp Ile Tyr Ala Lys Gly Cys Cys				
	260	265	270		
	Thr Ser Pro Glu Asp Gly Leu Glu Lys Ala Lys Lys Ile Gly Phe Pro				
	275	280	285		
	Val Met Ile Lys Ala Ser Glu Gly Gly Gly Gly Lys Gly Ile Arg Lys				
25	290	295	300		
	Val Asp Asp Glu Lys Asn Phe Ile Thr Leu Tyr Asn Gln Ala Ala Asn				
	305	310	315		320
	Glu Ile Pro Gly Ser Pro Ile Phe Ile Met Lys Leu Ala Gly Asp Ala				
	325	330	335		
30	Arg His Leu Glu Val Gln Leu Leu Ala Asp Gln Tyr Gly Thr Asn Ile				
	340	345	350		
	Ser Leu Phe Gly Arg Asp Cys Ser Val Gln Arg Arg His Gln Lys Ile				
	355	360	365		
	Ile Glu Glu Ala Pro Val Thr Ile Ala Arg Lys Glu Thr Phe His Glu				
35	370	375	380		
	Met Glu Asn Ala Ala Val Arg Leu Gly Lys Leu Val Gly Tyr Val Ser				
	385	390	395		400
	Ala Gly Thr Val Glu Tyr Leu Tyr Ser His Ala Glu Asp Lys Phe Tyr				
	405	410	415		
40	Phe Leu Glu Leu Asn Pro Arg Leu Gln Val Glu His Pro Thr Thr Glu				
	420	425	430		
	Met Val Thr Gly Val Asn Leu Pro Ala Ala Gln Leu Gln Ile Ala Met				
	435	440	445		
	Gly Ile Pro Met His Arg Ile Arg Asp Ile Arg Thr Leu Tyr Gly Ala				
45	450	455	460		
	Asp Pro His Thr Thr Thr Asp Ile Asp Phe Glu Phe Lys Ser Glu Thr				
	465	470	475		480

Ser Leu Val Ser Gln Arg Arg Pro Thr Pro Lys Gly His Cys Thr Ala
 485 490 495
 Cys Arg Ile Thr Ser Glu Asp Pro Gly Glu Gly Phe Lys Pro Ser Gly
 500 505 510
 5 Gly Ser Leu His Glu Leu Asn Phe Arg Ser Ser Ser Asn Val Trp Gly
 515 520 525
 Tyr Phe Ser Val Gly Asn Gln Ser Ser Ile His Ser Phe Ser Asp Ser
 530 535 540
 10 Gln Phe Gly His Ile Phe Ala Phe Gly Glu Asn Arg Gln Ala Ser Arg
 545 550 555 560
 Lys His Met Val Val Ala Leu Lys Glu Leu Ser Ile Arg Gly Asp Phe
 565 570 575
 Arg Thr Thr Val Glu Tyr Leu Ile Lys Leu Leu Glu Thr Pro Asp Phe
 580 585 590
 15 Glu Asp Asn Thr Ile Thr Thr Gly Trp Leu Asp Glu Leu Ile Thr Lys
 595 600 605
 Lys Leu Thr Ala Glu Arg Pro Asp Pro Ile Val Ala Val Val Cys Gly
 610 615 620
 20 Ala Val Thr Lys Ala His Ile Gln Ala Glu Glu Glu Lys Lys Glu Tyr
 625 630 635 640
 Ile Gln Ser Leu Glu Lys Gly Gln Val Pro His Arg Asn Leu Leu Lys
 645 650 655
 Thr Ile Phe Pro Val Glu Phe Ile Tyr Glu Gly Glu Arg Tyr Lys Phe
 660 665 670
 25 Thr Ala Thr Lys Ser Ser Glu Asp Lys Tyr Thr Leu Phe Leu Asn Gly
 675 680 685
 Ser Arg Cys Val Val Gly Ala Arg Ser Leu Ser Asp Gly Gly Leu Leu
 690 695 700
 30 Cys Ala Leu Asp Gly Lys Ser His Ser Val Tyr Trp Lys Glu Glu Ala
 705 710 715 720
 Ser Ala Thr Arg Leu Ser Val Asp Gly Lys Thr Cys Leu Leu Glu Val
 725 730 735
 Glu Asn Asp Pro Thr Gln Leu Arg Thr Pro Ser Pro Gly Lys Leu Val
 740 745 750
 35 Lys Tyr Leu Val Asp Ser Gly Glu His Val Asp Ala Gly Gln Pro Tyr
 755 760 765
 Ala Glu Val Glu Val Met Lys Met Cys Met Pro Leu Ile Ala Gln Glu
 770 775 780
 40 Asn Gly Val Val Gln Leu Ile Lys Gln Pro Gly Ser Thr Val Asn Ala
 785 790 795 800
 Gly Asp Ile Leu Ala Ile Leu Ala Leu Asp Asp Pro Ser Lys Val Lys
 805 810 815
 His Ala Lys Pro Phe Glu Gly Thr Leu Pro Ser Met Gly Glu Pro Asn
 820 825 830
 45 Val Thr Gly Thr Lys Pro Ala His Lys Phe Asn His Cys Ala Gly Ile
 835 840 845
 Leu Lys Asn Ile Leu Ala Gly Tyr Asp Asn Gln Val Ile Leu Asn Ser

850 855 860
 Thr Leu Lys Ser Leu Gly Glu Val Leu Lys Asp Asn Glu Leu Pro Tyr
 865 870 875 880
 Ser Glu Trp Gln Gln Ile Ser Ala Leu His Ser Arg Leu Pro Pro
 5 885 890 895
 Lys Leu Asp Asp Gly Leu Thr Ala Leu Val Glu Arg Thr Gln Ser Arg
 900 905 910
 Gly Ala Glu Phe Pro Ala Arg Gln Ile Leu Lys Leu Ile Thr Lys Ser
 915 920 925
 10 Ile Ala Glu Asn Gly Asn Asp Met Leu Glu Asp Val Val Ala Pro Leu
 930 935 940
 Val Ser Ile Ala Thr Ser Tyr Gln Asn Gly Leu Val Glu His Glu Tyr
 945 950 955 960
 15 Asp Tyr Phe Ala Ser Leu Ile Asn Glu Tyr Tyr Asp Val Glu Ser Leu
 965 970 975
 Phe Ser Gly Glu Asn Val Arg Glu Asp Asn Val Ile Leu Lys Leu Arg
 980 985 990
 Asp Glu Asn Lys Ser Asp Leu Lys Lys Val Ile Gly Ile Gly Leu Ser
 995 1000 1005
 20 His Ser Arg Val Ser Ala Lys Asn Asn Leu Ile Leu Ala Ile Leu Asp
 1010 1015 1020
 Ile Tyr Glu Pro Leu Leu Gln Ser Asn Ser Ser Val Ala Ala Ser Ile
 1025 1030 1035 1040
 Arg Glu Ala Leu Lys Asn Leu Phe Ile Arg Pro Arg Ala Cys Ala Lys
 25 1045 1050 1055
 Val Ala Leu Lys Ala Arg Glu Ile Leu Ile Gln Cys Ser Leu Pro Ser
 1060 1065 1070
 Ile Lys Glu Arg Ser Asp Gln Leu Glu His Ile Leu Arg Ser Ser Val
 1075 1080 1085
 30 Val Gln Thr Ser Tyr Gly Glu Ile Phe Ala Lys His Arg Glu Pro Asn
 1090 1095 1100
 Leu Glu Ile Ile Arg Glu Val Val Asp Ser Lys His Ile Val Phe Asp
 1105 1110 1115 1120
 Val Leu Ala Gln Phe Leu Ile Asn Pro Asp Pro Trp Val Ala Ile Ala
 35 1125 1130 1135
 Ala Ala Glu Val Tyr Val Arg Arg Ser Tyr Arg Ala Tyr Asp Leu Gly
 1140 1145 1150
 Lys Ile Glu Tyr His Val Asn Asp Arg Leu Pro Ile Val Glu Trp Lys
 1155 1160 1165
 40 Phe Lys Leu Ala Asn Met Gly Ala Ala Gly Val Asn Asp Ala Gln Gln
 1170 1175 1180
 Ala Ala Ala Ala Gly Gly Asp Asp Ser Thr Ser Met Lys His Ala Ala
 1185 1190 1195 1200
 Ser Val Ser Asp Leu Thr Phe Val Val Asp Ser Lys Thr Glu His Ser
 45 1205 1210 1215
 Thr Arg Thr Gly Val Leu Ala Pro Ala Arg His Leu Asp Asp Val Asp
 1220 1225 1230

-9-

Glu Thr Leu Thr Ala Ala Leu Glu Gln Phe Gln Pro Ala Asp Ala Ile
 1235 1240 1245
 Ser Phe Lys Ala Lys Gly Glu Thr Pro Glu Leu Leu Asn Val Leu Asn
 1250 1255 1260
 5 Ile Val Ile Thr Ser Ile Asp Gly Tyr Ser Asp Glu Asn Glu Tyr Leu
 1265 1270 1275 1280
 Ser Arg Ile Asn Glu Ile Leu Cys Glu Tyr Lys Glu Glu Leu Ile Ser
 1285 1290 1295
 10 Ala Gly Val Arg Arg Val Thr Phe Val Phe Ala His Gln Ile Gly Gln
 1300 1305 1310
 Tyr Pro Lys Tyr Tyr Thr Phe Thr Gly Pro Asp Tyr Glu Glu Asn Lys
 1315 1320 1325
 Val Ile Arg His Ile Glu Pro Ala Leu Ala Phe Gln Leu Glu Leu Gly
 1330 1335 1340
 15 Arg Leu Ala Asn Phe Asp Ile Lys Pro Ile Phe Thr Asn Asn Arg Asn
 1345 1350 1355 1360
 Ile His Val Tyr Asp Ala Ile Gly Lys Asn Ala Pro Ser Asp Lys Arg
 1365 1370 1375
 20 Phe Phe Thr Arg Gly Ile Ile Arg Thr Gly Val Leu Lys Glu Asp Ile
 1380 1385 1390
 Ser Ile Ser Glu Tyr Leu Ile Ala Glu Ser Asn Arg Leu Met Asn Asp
 1395 1400 1405
 Ile Leu Asp Thr Leu Glu Val Ile Asp Thr Ser Asn Ser Asp Leu Asn
 1410 1415 1420
 25 His Ile Phe Ile Asn Phe Ser Asn Ala Phe Asn Val Gln Ala Ser Asp
 1425 1430 1435 1440
 Val Glu Ala Ala Phe Gly Ser Phe Leu Glu Arg Phe Gly Arg Arg Leu
 1445 1450 1455
 30 Trp Arg Leu Arg Val Thr Gly Ala Glu Ile Arg Ile Val Cys Thr Asp
 1460 1465 1470
 Pro Gln Gly Thr Ser Phe Pro Leu Arg Ala Ile Ile Asn Asn Val Ser
 1475 1480 1485
 Gly Tyr Val Val Lys Ser Glu Leu Tyr Leu Glu Val Lys Asn Pro Lys
 1490 1495 1500
 35 Gly Glu Trp Val Phe Lys Ser Ile Gly His Pro Gly Ser Met His Leu
 1505 1510 1515 1520
 Arg Pro Ile Ser Thr Pro Tyr Pro Val Lys Glu Ser Leu Gln Pro Lys
 1525 1530 1535
 40 Arg Tyr Lys Ala His Asn Met Gly Thr Thr Tyr Val Tyr Asp Phe Pro
 1540 1545 1550
 Glu Leu Phe Arg Gln Ala Thr Ile Ser Gln Trp Lys Lys Tyr Gly Lys
 1555 1560 1565
 Lys Val Pro Lys Asp Val Phe Val Ser Leu Glu Leu Ile Thr Asp Glu
 1570 1575 1580
 45 Thr Asp Ser Leu Ile Ala Val Glu Arg Asp Pro Gly Ala Asn Lys Ile
 1585 1590 1595 1600
 Gly Met Val Gly Phe Lys Val Thr Ala Lys Thr Pro Glu Tyr Pro His

-10-

1605 1610 1615
 Gly Arg Gln Leu Ile Ile Val Ala Asn Asp Ile Thr His Lys Ile Gly
 1620 1625 1630
 5 Ser Phe Gly Pro Glu Glu Asp Asn Tyr Phe Asn Lys Cys Thr Glu Leu
 1635 1640 1645
 Ala Arg Lys Leu Gly Ile Pro Arg Ile Tyr Leu Ser Ala Asn Ser Gly
 1650 1655 1660
 Ala Arg Ile Gly Val Ala Glu Glu Leu Ile Pro Leu Tyr Gln Val Ala
 1665 1670 1675 1680
 10 Trp Asn Glu Glu Gly Ser Pro Asp Lys Gly Phe Arg Tyr Leu Tyr Leu
 1685 1690 1695
 Ser Thr Ala Ala Lys Glu Ser Leu Glu Lys Asp Gly Lys Ser Asp Ser
 1700 1705 1710
 Val Val Thr Glu Arg Ile Val Glu Lys Gly Glu Glu Arg His Val Ile
 15 1715 1720 1725
 Lys Ala Ile Ile Gly Ala Glu Asp Gly Leu Gly Val Glu Cys Leu Lys
 1730 1735 1740
 Gly Ser Gly Leu Ile Ala Gly Ala Thr Ser Arg Ala Tyr Lys Asp Ile
 1745 1750 1755 1760
 20 Phe Thr Ile Thr Leu Val Thr Cys Arg Ser Val Gly Ile Gly Ala Tyr
 1765 1770 1775
 Leu Val Arg Leu Gly Gln Arg Ala Ile Gln Ile Asp Gly Gln Pro Ile
 1780 1785 1790
 Ile Leu Thr Gly Ala Pro Ala Ile Asn Lys Leu Leu Gly Arg Glu Val
 25 1795 1800 1805
 Tyr Ser Ser Asn Leu Gln Leu Gly Gly Thr Gln Ile Met Tyr Asn Asn
 1810 1815 1820
 Gly Val Ser His Leu Thr Ala Asn Asp Asp Leu Ala Gly Val Glu Lys
 1825 1830 1835 1840
 30 Ile Met Glu Trp Leu Ser Tyr Val Pro Ala Lys Arg Gly Leu Pro Val
 1845 1850 1855
 Pro Ile Leu Glu Ser Glu Asp Ser Trp Asp Arg Asp Val Asp Tyr Tyr
 1860 1865 1870
 Pro Pro Lys Gln Glu Ala Phe Asp Val Arg Trp Met Ile Gln Gly Arg
 35 1875 1880 1885
 Glu Val Asp Gly Glu Tyr Glu Ser Gly Leu Phe Asp Lys Asp Ser Phe
 1890 1895 1900
 Gln Glu Thr Leu Ser Gly Trp Ala Lys Gly Val Val Val Gly Arg Ala
 1905 1910 1915 1920
 40 Arg Leu Gly Gly Ile Pro Ile Gly Val Ile Gly Val Glu Thr Arg Thr
 1925 1930 1935
 Val Glu Asn Leu Ile Pro Ala Asp Pro Ala Asn Pro Asp Ser Thr Glu
 1940 1945 1950
 Ser Leu Ile Gln Glu Ala Gly Gln Val Trp Tyr Pro Asn Ser Ala Phe
 45 1955 1960 1965
 Lys Thr Ala Gln Ala Ile Asn Asp Phe Asn Asn Gly Glu Gln Leu Pro
 1970 1975 1980

-11-

Leu Met Ile Leu Ala Asn Trp Arg Gly Phe Ser Gly Gly Gln Arg Asp
 1985 1990 1995 2000
 Met Tyr Asn Glu Val Leu Lys Tyr Gly Ser Phe Ile Val Asp Ala Leu
 2005 2010 2015
 5 Val Asp Phe Lys Gln Pro Ile Phe Thr Tyr Ile Pro Pro Asn Gly Glu
 2020 2025 2030
 Leu Arg Gly Gly Ser Trp Val Val Val Asp Pro Thr Ile Asn Ser Asp
 2035 2040 2045
 10 Met Met Glu Met Tyr Ala Asp Val Asp Ser Arg Ala Gly Val Leu Glu
 2050 2055 2060
 Pro Glu Gly Met Val Gly Ile Lys Tyr Arg Arg Asp Lys Leu Leu Ala
 2065 2070 2075 2080
 Thr Met Glu Arg Leu Asp Pro Thr Tyr Gly Glu Met Lys Ala Lys Leu
 2085 2090 2095
 15 Asn Asp Ser Ser Leu Ser Pro Glu Glu His Ser Lys Ile Ser Ala Lys
 2100 2105 2110
 Leu Phe Ala Arg Glu Lys Ala Leu Leu Pro Ile Tyr Ala Gln Ile Ser
 2115 2120 2125
 20 Val Gln Phe Ala Asp Leu His Asp Arg Ser Gly Arg Met Leu Ala Lys
 2130 2135 2140
 Gly Val Ile Arg Lys Glu Ile Lys Trp Thr Asp Ala Arg Arg Phe Phe
 2145 2150 2155 2160
 Phe Trp Arg Leu Arg Arg Arg Leu Asn Glu Glu Tyr Val Leu Arg Leu
 2165 2170 2175
 25 Ile Ser Glu Gln Ile Lys Asp Ser Ser Lys Leu Glu Arg Val Ala Arg
 2180 2185 2190
 Leu Lys Ser Trp Met Pro Thr Val Glu Tyr Asp Asp Asp Gln Ala Val
 2195 2200 2205
 30 Ser Asn Trp Ile Glu Glu Asn His Ala Lys Leu Gln Lys Arg Val Asn
 2210 2215 2220
 Glu Leu Lys Gln Glu Val Ser Arg Thr Lys Ile Met Arg Leu Leu Lys
 2225 2230 2235 2240
 Glu Asp Pro Asn Ser Ala Ile Ser Ala Met Lys Asp Tyr Val Glu Arg
 2245 2250 2255
 35 Leu Ser Lys Glu Asp Lys Glu Lys Phe Leu Lys Ala Leu Lys
 2260 2265 2270

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/GB 98/03857

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/52 C12N9/00 C07K16/40 C12N15/11 C12N15/81
C12Q1/25

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FR 2 727 129 A (RHONE POULENC AGROCHIMIE) 24 May 1996	1-13, 16, 17
Y	see the whole document ---	14, 15
X	AL-FEEL W ET AL: "Cloning of the yeast FAS3 gene and primary structure of yeast acetyl-CoA carboxylase" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 89, May 1992, pages 4534-4538, XP002097900 WASHINGTON US see the whole document ---	1-12
Y	GB 2 137 208 A (COLLABORATIVE RES INC) 3 October 1984 see the whole document ---	14, 15
-/--		

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

25 March 1999

Date of mailing of the international search report

09/04/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Van der Schaal, C

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 98/03857

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>HORIKAWA S ET AL: "CELL-FREE TRANSLATION AND REGULATION OF CANDIDA -LIPOLYTICA ACETYL COENZYME A CARBOXYLASE EC-6.4:1.2 MESSENGER RNA." EUR J BIOCHEM, (1980) 104 (1), 191-198. CODEN: EJBCAI. ISSN: 0014-2956., XP002097901</p> <p>-----</p>	1-12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 98/03857

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
FR 2727129 A	24-05-1996	NONE	
GB 2137208 A	03-10-1984	US 4661454 A	28-04-1987
		AT 64416 T	15-06-1991
		CA 1283373 A	23-04-1991
		CA 1273883 C	11-09-1990
		DK 97784 A	29-08-1984
		EP 0123811 A	07-11-1984
		JP 60058077 A	04-04-1985
		US 5139936 A	18-08-1992

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)